

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2003, 08:45:11 ; Search time 38 Seconds

(without alignments)
1413.158 Million cell updates/sec

Title: US-09-412-100-23

Perfect score: 2079

Sequence: 1 MSMTSGICASTMQISIGCA.....DAMAGDAINNALGKLGAA 403

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2079	100.0	403	19	AAW75863
2	2079	100.0	403	19	AAW62455
3	2079	100.0	403	19	AAW61114
4	2079	100.0	403	20	AAW87639
5	2079	100.0	403	21	AAW71093
6	2079	100.0	403	21	AAW84854
7	2079	100.0	403	22	AAW67710
8	2079	100.0	403	22	ABW09225
9	2079	100.0	403	23	AAE18295
10	2079	100.0	403	23	AAE16447

11	1928	92.7	385	18	AAW06598
12	1913	92.0	385	15	AAW45751
13	718.5	34.6	338	18	AAW06597
14	718.5	34.6	338	20	AAW87638
15	718.5	34.6	338	20	AAW82407
16	718.5	34.6	338	21	AAW71092
17	718.5	34.6	338	21	AAW84853
18	718.5	34.6	338	21	AAW5801
19	718.5	34.6	338	22	AAW06709
20	718.5	34.6	338	23	AAW09224
21	718.5	34.6	338	23	AAE18294
22	718.5	34.6	338	23	AAE16446
23	718.5	34.6	340	19	AAW75862
24	718.5	34.6	340	19	AAW61113
25	713.5	34.3	340	19	AAW62454
26	226.5	10.9	898	18	AAW31853
27	222	10.7	272	22	ABW70198
28	215.5	10.4	718	12	AAW14308
29	211.5	10.2	651	20	AAW40097
30	211.5	10.2	651	23	AAW11781
31	211.5	10.2	718	19	AAW53346
32	211.5	10.2	718	21	AAW53070
33	206	9.9	604	16	AAW99057
34	204	9.8	528	22	AAW82611
35	203.5	9.8	318	21	AAW81229
36	202	9.7	630	23	AAW50042
37	202	9.7	676	23	AAW50047
38	202	9.7	989	23	AAW50038
39	202	9.7	1255	23	AAW50037
40	202	9.7	1880	23	AAW50039
41	200	9.6	388	20	AAW04999
42	199.5	9.6	201	21	AAW31750
43	198.5	9.5	738	19	AAW56163
44	198.5	9.5	3190	22	AAW84634
45	198.5	9.5	3275	22	AAW70437

ALIGNMENTS

RESULT 1
AAW75863
ID AAW75863 standard; Protein; 403 AA.
XX
AC AAW75863;
XX
DT 07-DEC-1998 (first entry)
XX
DE Erwinia amylovora hypersensitive response elicitor (HRE).
XX
KW Hypersensitive response elicitor; HRE; insect resistance;
XX biological control; transgenic plant.
XX
OS Erwinia amylovora.
XX
PN W09837752-A1.
XX
PD 03-SEP-1998.
XX
PF 26-FEB-1998; 98WO-US03604.
XX
PR 28-FEB-1997; 97US-0039226.
XX
(CORR) CORNELL RES FOUND INC.
PA Wel Z, Zitter TA;
XX
XX WPI: 1998-495374/42.
XX N-PSDB; AAW54607.
XX
XX Use of hypersensitive response elicitor polypeptide - for
XX application to plants or seeds or transgenic plants or seeds for the
XX control of insects.
PT

XX Disclosure; Page 9-10; 75pp; English.

CC This is the amino acid sequence of a 39 kDa, heat stable
CC hypersensitive response elicitor (HRE) of *Erwinia amylovora*. The
CC invention relates to the use of a HRE polypeptide or protein to
CC control insects on plants or plants grown from seed treated with HRE.
CC Also claimed is a method of insect control for plants that involves:
CC (a) providing a transgenic plant or seed transformed with a DNA
CC molecule (see AAV54606-09) encoding a HRE polypeptide or protein (see
CC AAV5862-67); and (b) growing the transgenic plants or transgenic
CC plants produced from the transgenic seeds to control insects. HRE
CC prevents direct insect damage to plants by feeding injury. It kills
CC insects close to plants, and interferes with insect larval feeding
CC on such plants. It also prevents insects from colonizing host
CC plants and releasing phytotoxins which result in disease damage to
CC plants.

CC Sequence 403 AA;

Query Match 100.0%; Score 2079; DB 19; Length 403;
Best Local Similarity 100.0%; Pred. No. 3.2e-155;
Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSINTSGIGASTMJOISIGAGGNNGLGTSRONAGIGGNSALGIGGNNNDVYNOLAGLL 60
DB 1 MSINTSGIGASTMJOISIGAGGNNGLGTSRONAGIGGNSALGIGGNNNDVYNOLAGLL 60
QY 61 TGMAMMMAMMGCGGGLMGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGL 120
DB 61 TGMAMMMAMMGCGGGLMGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGL 120
QY 121 TTSTNSPLDQALGINSTSONDSTSGTSTSDSSDPMOQLKMFSEIMQSLFGDGQDGT 180
DB 121 TTSTNSPLDQALGINSTSONDSTSGTSTSDSSDPMOQLKMFSEIMQSLFGDGQDGT 180
QY 181 QGSSSGGKOPTGEONAYRKGVTDALSGIMNGLSQLLNGGLGGLGGLGGLGGLGGLGGLGGL 240
DB 181 QGSSSGGKOPTGEONAYRKGVTDALSGIMNGLSQLLNGGLGGLGGLGGLGGLGGLGGLGGL 240
QY 241 GKGGLQNLGSPVDYQOLGNAVGTGIMKAGIQALNDIGTHRSSSTFSFVNKGDRAAKEI 300
DB 241 GKGGLQNLGSPVDYQOLGNAVGTGIMKAGIQALNDIGTHRSSSTFSFVNKGDRAAKEI 300
QY 301 GGFMDQYEVFEGKPYOKPGQEVKTDDSKMAKALSKPDDGMPASMEQFNKAGMIR 360
DB 301 GGFMDQYEVFEGKPYOKPGQEVKTDDSKMAKALSKPDDGMPASMEQFNKAGMIR 360
QY 361 PMAGDTGNGNLOARAGAGSSSLGIDAMMAGDAINNNALGRLGAA 403
DB 361 PMAGDTGNGNLOARAGAGSSSLGIDAMMAGDAINNNALGRLGAA 403

RESULT 2
AAW62455
ID AAW62455 standard; Protein; 403 AA.

AC AAW62455;

DT 09-NOV-1998 (first entry)

DE *Erwinia amylovora* hypersensitive response elicitor (HRE).

FM Hypersensitive response elicitor; HRE; growth; transgenic plant.

OS *Erwinia amylovora*.

PN WO9832844-A1.

PD 30-JUL-1998.

PF 27-JAN-1998; 98WO-US01507.

PR 27-JAN-1997; 97US-0036048.
PA (CORR) CORNELL RES FOUND INC.
PI Beer SV, Qiu D, Wei Z;
XX

DR WPI, 1998-427940/36.
DR N-PDB; AAV39973.

PT Method for enhancing plant growth - comprises use of hypersensitive
PT response elicitor polypeptide or protein which may also effect, e.g.
PT increase in plant height or earlier germination seed
PS Disclosure; Page 15-16; 110pp; English.

CC This is the deduced amino acid sequence of the 39 kDa, heat stable
CC hypersensitive response elicitor (HRE) of *Erwinia amylovora*. A
CC method of enhancing growth in plants comprises: (a) applying a HRE
CC polypeptide or protein in a non-infectious form to a plant or plant
CC seed under conditions effective to enhance growth of the plant or
CC plants grown from the seed, or (b) providing a transgenic plant or
CC plant seed transformed with a DNA molecule encoding a HRE
CC polypeptide or protein, and growing the transgenic plant or a plant
CC produced from the transgenic seed under conditions effective to
CC enhance plant growth. HRES (see AAW62454-59) or nucleic acids
CC encoding them (see AAV39972-75) can be used to increase plant growth.
CC The HRES may also result in increased plant height and yield, and
CC effect early germination and maturation of plant seed and early
CC coloration of fruit and plants. *E. amylovora* HRE can be applied
CC to tomato plants to enhance growth without causing disease in that
CC species; this bacterium is a pathogen of apple and pear but not
CC of tomato.

CC Sequence 403 AA;

Query Match 100.0%; Score 2079; DB 19; Length 403;
Best Local Similarity 100.0%; Pred. No. 3.2e-155;
Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSINTSGIGASTMJOISIGAGGNNGLGTSRONAGIGGNSALGIGGNNNDVYNOLAGLL 60
DB 1 MSINTSGIGASTMJOISIGAGGNNGLGTSRONAGIGGNSALGIGGNNNDVYNOLAGLL 60
QY 61 TGMAMMMAMMGCGGGLMGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGL 120
DB 61 TGMAMMMAMMGCGGGLMGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGL 120
QY 121 TTSTNSPLDQALGINSTSONDSTSGTSTSDSSDPMOQLKMFSEIMQSLFGDGQDGT 180
DB 121 TTSTNSPLDQALGINSTSONDSTSGTSTSDSSDPMOQLKMFSEIMQSLFGDGQDGT 180
QY 181 QGSSSGGKOPTGEONAYRKGVTDALSGIMNGLSQLLNGGLGGLGGLGGLGGLGGLGGLGGL 240
DB 181 QGSSSGGKOPTGEONAYRKGVTDALSGIMNGLSQLLNGGLGGLGGLGGLGGLGGLGGLGGL 240
QY 241 GKGGLQNLGSPVDYQOLGNAVGTGIMKAGIQALNDIGTHRSSSTFSFVNKGDRAAKEI 300
DB 241 GKGGLQNLGSPVDYQOLGNAVGTGIMKAGIQALNDIGTHRSSSTFSFVNKGDRAAKEI 300
QY 301 GGFMDQYEVFEGKPYOKPGQEVKTDDSKMAKALSKPDDGMPASMEQFNKAGMIR 360
DB 301 GGFMDQYEVFEGKPYOKPGQEVKTDDSKMAKALSKPDDGMPASMEQFNKAGMIR 360
QY 361 PMAGDTGNGNLOARAGAGSSSLGIDAMMAGDAINNNALGRLGAA 403
DB 361 PMAGDTGNGNLOARAGAGSSSLGIDAMMAGDAINNNALGRLGAA 403

RESULT 3
AAW61114

ID AAW61114 standard; Protein; 403 AA.

AC AAW61114;

```

XX 26-OCT-1998 (first entry)
DE Hypersensitive response elicitor protein (39 kDa).
KW Hypersensitive response elicitor; transgenic plant; seed;
XX pathogen resistance; disease resistance; crop protection.
OS Erwinia amylovora.
XX MO9824297-A1.
XX 11-JUN-1998.
XX 04-DEC-1997; 97MO-US22629.
XX 05-DEC-1996; 96DS-0033230.
XX (CORR ) CORNELL RES FOUND INC.
XX Beer SV, Qiu D, Wei Z;
XX WPI; 1998-332931/29.
XX N-PSDB; AAV36428.
XX Imparting pathogen resistance to plants - by applying a
XX hypersensitive response elicitor polypeptide to seeds
XX Disclosure; Page 18-20; 85pp; English.
XX This is the 39 kDa hypersensitive elicitor (HRE) protein of
XX Erwinia amylovora. It is heat stable at 100 degC for at least 10
XX min, has a pI of approximately 4.3, and contains substantially no
XX cysteine. The invention relates to methods of imparting no
XX hypersensitive response induced resistance to plants by treatment
XX of seeds. Isolated HRE proteins can be applied to seeds as a means
XX of imparting pathogen resistance to plants grown from the seeds.
XX Alternatively, bacteria containing the gene encoding the HRE can be
XX applied to the plant seeds, or transgenic plant seeds containing a
XX DNA molecule encoding an HRE polypeptide or protein are used. HRE
XX polypeptide sequences from Erwinia chrysanthemi, Erwinia amylovora,
XX Pseudomonas syringae, Pseudomonas solanacearum, Xanthomonas
XX campestris pv. glycines and Xanthomonas campestris pelargonii (see
XX AAM6113-18) are provided. The methods can impart pathogen
XX resistance without using agents which are harmful to the
XX environment or pathogenic to the plant seed being treated, or to
XX adjacent plants. E. amylovora causes disease in apple or pear
XX but not tomato. However, it elicits a hypersensitive response in
XX tomato. Thus, E. amylovora can be applied to tomato seeds to
XX impart pathogen resistance without causing diseases in plants of
XX that species.
XX
XX Sequence 403 AA:
XX
XX Query Match 100.0%; Score 2079; DB 19; Length 403;
XX Best Local Similarity 100.0%; Pred. No. 3.2e-155;
XX Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 241 GGGKGLMTSGPVDOOLGNVGTGKAGAGQALNDIGTRHSTRSFVNRKGDRAAKEI 300
DB 241 GGGKGLMTSGPVDOOLGNVGTGKAGAGQALNDIGTRHSTRSFVNRKGDRAAKEI 300
OY 301 GGFMDQYPEVFGKRPQYOKPGQEVKTDKSWAKALSRPDDGKMPASMEQFNKAKMIKR 360
DB 301 GGFMDQYPEVFGKRPQYOKPGQEVKTDKSWAKALSRPDDGKMPASMEQFNKAKMIKR 360
OY 361 PMAGDTGNGNLQARAGAGSSSLGIDAMAGDAINNMALGKLGAA 403
DB 361 PMAGDTGNGNLQARAGAGSSSLGIDAMAGDAINNMALGKLGAA 403
RESULT 4
AAM87639
ID AAM87639 standard; Protein; 403 AA.
XX
XX AAM87639;
XX
XX 09-MAR-1999 (first entry)
XX
XX A hypersensitive response elicitor protein.
XX
XX Hypersensitive response elicitor protein; hairpin protein;
XX disease resistance; seed quality; insect control; corn borer;
XX Lepidoptera larvae; transgenic plant.
XX
XX Erwinia amylovora.
XX
XX MO9854214-A2.
XX
XX 03-DEC-1998.
XX
XX 28-MAY-1998; 98MO-US10874.
XX
XX 30-MAY-1997; 97US-0048109.
XX
XX (CORR ) CORNELL RES FOUND INC.
XX (EDEN-) EDEN BIOSCIENCE CORP.
XX
XX Beer SV, Laby RJ, Wei Z;
XX
XX WPI; 1999-070210/06.
XX N-PSDB; AAV83989.
XX
XX New fragments of an Erwinia hypersensitive response elicitor protein
XX and related DNA - used to impart disease resistance to plants, to
XX increase their growth and to control insects
XX
XX Claim 4; Page 10-11; 94pp; English.
XX
XX The present sequence represents a hypersensitive response elicitor
XX protein (also called hairpin protein) that is able to elicit a
XX hypersensitive response in plants. The specification also describes
XX hypersensitive response elicitors from other pathogenic organisms.
XX The protein, in non-infectious form, is applied to plants to impart
XX disease resistance (to a wide range of viral, bacterial and fungal
XX pathogens), to improve growth (yield, quantity and quality of seeds,
XX to provide earlier germination etc.) and to control insects (e.g. corn
XX borers, Lepidoptera larvae etc.) The same results are provided by
XX transgenic plants expressing the protein.
XX
XX Sequence 403 AA:
XX
XX Query Match 100.0%; Score 2079; DB 20; Length 403;
XX Best Local Similarity 100.0%; Pred. No. 3.2e-155;
XX Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 61 TGMAMMSMGGGGLMGGGLGGLGNGIGSSGGLGEGLSNALNDMLGSLNTLGSKGGNN 120
 DB 61 TGMAMMSMGGGGLMGGGLGGLGNGIGSSGGLGEGLSNALNDMLGSLNTLGSKGGNN 120
 QY 121 TTSTNSPLDQALGINSTSONDSTSGTSDTSDDPMQOLLKMFSEIMQSLFEDGDGT 180
 DB 121 TTSTNSPLDQALGINSTSONDSTSGTSDTSDDPMQOLLKMFSEIMQSLFEDGDGT 180
 QY 181 QGSSSGKOPTREBQNAKRYKGYTDALSLGMLNGISQLLNGGGLGGGCGGNAAGTGLDSSSL 240
 DB 181 QGSSSGKOPTREBQNAKRYKGYTDALSLGMLNGISQLLNGGGLGGGCGGNAAGTGLDSSSL 240
 QY 241 GKGGLQNLGSPVDYQOLGNAVGTGIGMKAGIOALNDIGTRHSSTRSFYKNGDRAMAKEY 300
 DB 241 GKGGLQNLGSPVDYQOLGNAVGTGIGMKAGIOALNDIGTRHSSTRSFYKNGDRAMAKEY 300
 QY 301 GGFMDQYPEYFGKPOYOKGPGQEVKTDKSWAKALSRPDDGKMTPASMEQFNKAKGMIKR 360
 DB 301 GGFMDQYPEYFGKPOYOKGPGQEVKTDKSWAKALSRPDDGKMTPASMEQFNKAKGMIKR 360
 QY 361 PMAGDTGNGNLQARGAGSSSLGIDAMMAGDAINNMALGKLGAA 403
 DB 361 PMAGDTGNGNLQARGAGSSSLGIDAMMAGDAINNMALGKLGAA 403

RESULT 5

AA71093
 ID AA71093 standard; Protein: 403 AA.

AC AA71093;

DT 08-SEP-2000 (first entry)

DE Erwilia amylovora hypersensitive response elicitor #1.

KW Hypersensitive response elicitor; environmental stress resistance;
 plant.

OS Erwilia amylovora.

PN W0200028055-A2.

PD 18-MAY-2000.

PF 04-NOV-1999; 99WO-US26039.

PR 05-NOV-1998; 98US-0107243.

PA (EDEN-) EDEN BIOSCIENCE CORP.

PI Wei Z, Schading RL;

DR WPI, 2000-376566/32.

DR N-PSDB; AAD00668.

XX Application of a hypersensitive response elicitor protein to plants to
 PT impart stress resistance

PS Disclosure; Page 7-8; 84pp; English.

CC The patent discloses a method to impart stress resistance to plants by
 CC applying a hypersensitive response elicitor in a non-infectious form to
 CC a plant or seed. The present sequence is a hypersensitive
 CC response elicitor protein from Erwilia amylovora. It is
 CC is used to impart stress resistance to plants.

XX Sequence 403 AA;

QY Query Match 100.0%; Score 2079; DB 21; Length 403;

Best Local Similarity 100.0%; Pred. No. 3.2e-153;
 Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSINTSGIAGASTMGIISIGAGGNNGLGTSRQNAAGLGNALGIGGNGNDYVQNLGL 60

DB 1 MSINTSGIAGASTMGIISIGAGGNNGLGTSRQNAAGLGNALGIGGNGNDYVQNLGL 60
 QY 61 TGMAMMSMGGGGLMGGGLGGLGNGIGSSGGLGEGLSNALNDMLGSLNTLGSKGGNN 120
 DB 61 TGMAMMSMGGGGLMGGGLGGLGNGIGSSGGLGEGLSNALNDMLGSLNTLGSKGGNN 120
 QY 121 TTSTNSPLDQALGINSTSONDSTSGTSDTSDDPMQOLLKMFSEIMQSLFEDGDGT 180
 DB 121 TTSTNSPLDQALGINSTSONDSTSGTSDTSDDPMQOLLKMFSEIMQSLFEDGDGT 180
 QY 181 QGSSSGKOPTREBQNAKRYKGYTDALSLGMLNGISQLLNGGGLGGGCGGNAAGTGLDSSSL 240
 DB 181 QGSSSGKOPTREBQNAKRYKGYTDALSLGMLNGISQLLNGGGLGGGCGGNAAGTGLDSSSL 240
 QY 241 GKGGLQNLGSPVDYQOLGNAVGTGIGMKAGIOALNDIGTRHSSTRSFYKNGDRAMAKEY 300
 DB 241 GKGGLQNLGSPVDYQOLGNAVGTGIGMKAGIOALNDIGTRHSSTRSFYKNGDRAMAKEY 300
 QY 301 GGFMDQYPEYFGKPOYOKGPGQEVKTDKSWAKALSRPDDGKMTPASMEQFNKAKGMIKR 360
 DB 301 GGFMDQYPEYFGKPOYOKGPGQEVKTDKSWAKALSRPDDGKMTPASMEQFNKAKGMIKR 360
 QY 361 PMAGDTGNGNLQARGAGSSSLGIDAMMAGDAINNMALGKLGAA 403
 DB 361 PMAGDTGNGNLQARGAGSSSLGIDAMMAGDAINNMALGKLGAA 403

RESULT 6

AA84854
 ID AA84854 standard; Protein: 403 AA.

AC AA84854;

DT 08-AUG-2000 (first entry)

DE A hypersensitive response elicitor protein.

KW Hypersensitive response; insect control; disease resistance;
 KW hypersensitive response elicitor; plant growth; vegetable; crop;
 KW ornamental plant.

OS Erwilia amylovora.

XX Location/Qualifiers

FT Key 169..403

FT Peptide /note="claimed under claim 5"

FT Peptide 210..403

FT Peptide /note="claimed under claim 5"

FT Peptide 267..403

FT Peptide /note="claimed under claim 5"

FT Peptide 343..403

FT Peptide /note="claimed under claim 5"

FT Peptide 105..179

FT Peptide /note="claimed under claim 7"

FT Peptide 121..150

FT Peptide /note="claimed under claim 7"

FT Peptide 137..166

FT Peptide /note="claimed under claim 7"

FT Peptide 137..156

FT Peptide /note="claimed under claim 7"

PN W0200020452-A2.

PD 13-APR-2000.

PF 05-OCT-1999; 99WO-US23181.

PR 05-OCT-1998; 98US-0103050.

PA (EDEN-) EDEN BIOSCIENCE CORP.

PI Wei Z, Fan H, Niggemeyer JL;

QY 1 MSINTSGLASTMOWISIGAGAGNNGLGTSRONAGLGNSALGLGGGNDNDYVNOIAGL 60
DB 1 MSINTSGLASTMOWISIGAGAGNNGLGTSRONAGLGNSALGLGGGNDNDYVNOIAGL 60
QY 61 TGMAMMAMMMGGGGLMGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGL 120
DB 61 TGMAMMAMMMGGGGLMGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGL 120
QY 121 TTSTNSPLDQALGINSTSONDSTGCTDSTSDSPMQQLKMFSEIMOSLFGDGQGT 180
DB 121 TTSTNSPLDQALGINSTSONDSTGCTDSTSDSPMQQLKMFSEIMOSLFGDGQGT 180
QY 181 QGSSSGGKOPTEBEQONAKKGYTDALSGLMGNGLSQLLNGGGLGGGGGAGNAGTGLDSSL 240
DB 181 QGSSSGGKOPTEBEQONAKKGYTDALSGLMGNGLSQLLNGGGLGGGGGAGNAGTGLDSSL 240
QY 241 GGGKGLQNLGSPVDYQOLGNAVGTGIGMKAGIOLANDIGTHRSSTSFVKKGRAMAKET 300
DB 241 GGGKGLQNLGSPVDYQOLGNAVGTGIGMKAGIOLANDIGTHRSSTSFVKKGRAMAKET 300
QY 301 GGFMDQYPEVFGKPOYOKGPGOEVTDDKSWAKALSKPDDGWTSPASMEQFNKAKGMIKR 360
DB 301 GGFMDQYPEVFGKPOYOKGPGOEVTDDKSWAKALSKPDDGWTSPASMEQFNKAKGMIKR 360
QY 361 PMAGDTGNGNLOARGAGSSSLGIDAMMAGDAINNMAIGKLGAA 403
DB 361 PMAGDTGNGNLOARGAGSSSLGIDAMMAGDAINNMAIGKLGAA 403

RESULT 10
AAE16447
ID AAE16447 standard; Protein; 403 AA.
AC AAE16447;
XX
DT 09-APR-2002 (first entry)
XX
DE E. amylovora hypersensitive response elicitor protein #1.
XX
XX Hypersensitive response elicitor protein; plant growth; fruit coloration;
KW disease resistance; stress resistance; phytoalexin; insect infection;
KW plant maturation.
XX
OS Erwina amylovora.
XX
FH Key Location/Qualifiers
FT Domain 32..74
FT /label= Hypersensitive_response_eliciting_domain_1
FT Region 32..57
FT /label= Acidic_unit
FT Region 57..74
FT /label= Alpha_helix
FT Domain 130..180
FT /label= Hypersensitive_response_eliciting_domain_2
FT Region 130..157
FT /label= Acidic_unit
FT Region 157..180
FT /label= Alpha_helix
FT
FT WO200196501-A2.
XX
XX
XX 27-DEC-2001.
XX
XX 12-JUN-2001; 2001MO-US18820.
XX
XX 16-JUN-2000; 2000US-212211P.
XX
XX (EDEN-) EDEN BIOSCIENCE CORP.
XX
XX Pan H, Wei Z;
XX
XX WPI; 2002-122282/16.
XX
XX

DR N-PSDB; AAD27015.
XX
XX New hypersensitive response elicitor proteins comprising spaced apart
PT domains having an acidic portion linked to an alpha-helix, useful for
PT imparting disease or stress resistance, controlling insects or
PT enhancing plant growth -
XX
XX
PS Disclosure; Page 9-10; 9pp; English.
XX
XX The patent discloses hypersensitive response elicitor proteins and
CC nucleotides encoding such proteins. Hypersensitive response elicitor
CC proteins comprise an isolated pair or more of spaced apart domains,
CC each comprising an acidic portion linked to an alpha-helix and capable
CC of eliciting a hypersensitive response in plants. Sequences of the
CC invention are used to impart disease resistance to plants, to enhance
CC plant growth, to control insects and/or to impart stress resistance
CC to plants which includes resistance to environmental stresses such as
CC climate, air pollution, chemical and nutritional stress. The method of
CC imparting disease resistance has the potential for treating previously
CC untreatable diseases, treating diseases systemically and avoiding the
CC use of infectious agents or environmentally harmful materials. Hyper-
CC sensitive response elicitor sequences are used to enhance plant growth
CC which encompasses greater yield, increased in quantity of seeds produced,
CC percentage of seeds germinated, plant size and biomass, bigger fruits,
CC earlier fruit coloration and plant maturation. They are also used for
CC insect control which encompasses preventing direct insect damage to
CC plant by feeding injury, interfering with insect larval feeding on the
CC plants, preventing insects from colonizing host plants and releasing
CC phytochemicals. Sequences of the invention also prevent subsequent disease
CC damage to plants resulting from insect infection. The present sequence
CC is Erwina amylovora hypersensitive response elicitor protein.
XX
SQ Sequence 403 AA;
Query Match 100.0%; Score 2079; DB 23; Length 403;
Best Local Similarity 100.0%; Pred. No. 3.2e-155;
Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSINTSGLASTMOWISIGAGAGNNGLGTSRONAGLGNSALGLGGGNDNDYVNOIAGL 60
DB 1 MSINTSGLASTMOWISIGAGAGNNGLGTSRONAGLGNSALGLGGGNDNDYVNOIAGL 60
QY 61 TGMAMMAMMMGGGGLMGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGL 120
DB 61 TGMAMMAMMMGGGGLMGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGL 120
QY 121 TTSTNSPLDQALGINSTSONDSTGCTDSTSDSPMQQLKMFSEIMOSLFGDGQGT 180
DB 121 TTSTNSPLDQALGINSTSONDSTGCTDSTSDSPMQQLKMFSEIMOSLFGDGQGT 180
QY 181 QGSSSGGKOPTEBEQONAKKGYTDALSGLMGNGLSQLLNGGGLGGGGGAGNAGTGLDSSL 240
DB 181 QGSSSGGKOPTEBEQONAKKGYTDALSGLMGNGLSQLLNGGGLGGGGGAGNAGTGLDSSL 240
QY 241 GGGKGLQNLGSPVDYQOLGNAVGTGIGMKAGIOLANDIGTHRSSTSFVKKGRAMAKET 300
DB 241 GGGKGLQNLGSPVDYQOLGNAVGTGIGMKAGIOLANDIGTHRSSTSFVKKGRAMAKET 300
QY 301 GGFMDQYPEVFGKPOYOKGPGOEVTDDKSWAKALSKPDDGWTSPASMEQFNKAKGMIKR 360
DB 301 GGFMDQYPEVFGKPOYOKGPGOEVTDDKSWAKALSKPDDGWTSPASMEQFNKAKGMIKR 360
QY 361 PMAGDTGNGNLOARGAGSSSLGIDAMMAGDAINNMAIGKLGAA 403
DB 361 PMAGDTGNGNLOARGAGSSSLGIDAMMAGDAINNMAIGKLGAA 403

RESULT 11
AAW06598
ID AAW06598 standard; Protein; 385 AA.
XX
XX AAW06598;
XX
XX

DT	30-MAR-1997	(first entry)
XX		
DE		Hypersensitive response elicitor protein.
XX		
KM		Hypersensitive response; elicitor; <i>Erythrina amylovora</i> ; plant;
KM		disease-resistance; <i>Escherichia coli</i> ; infiltration; virus;
KM		bacterium; fungus; pathogen; biological control agent.
XX		
OS		<i>Erythrina amylovora</i> .
XX		
PN		MO9639802-AL.
XX		
PD		19-DEC-1996.
XX		
PF		05-JUN-1996; 96WO-US08819.
XX		
PR		07-JUN-1995; 95US-0475775.
XX		
PA		(CORR) CORNELL RES FOUND INC.
XX		
PI		Beer SV, Wei Z;
XX		
DR		WPI; 1997-051614/05.
DR		N-PESDB; AAF49314.
XX		
PT		Imparting pathogen resistance to plants - with hypersensitive
XX		response elicitor polypeptide or protein
PS		Claim 7; Page 46-47; 69pp; English.
XX		
CC		This sequence represents a hypersensitive response elicitor from
CC		<i>Erythrina amylovora</i> , with a mol.wt. of 37,000. The elicitor has a pI
CC		of 4.3, thermostability at 100 deg C for at least 10 min, and
CC		contains no cysteine. The elicitor may be used in a new method for
CC		impairing pathogen resistance to plants, by application of the
CC		elicitor in a non-infectious form to plant cells, by spraying,
CC		injection, leaf abrasion, or plant infection with recombinant
CC		bacteria (non-infectious to the host plant, e.g. <i>Escherichia coli</i>)
CC		expressing the elicitor as a biological control agent, to allow
CC		recombinant protein infiltration into the plant. The method
CC		confers virus, bacterium or fungus disease-resistance on crops and
CC		ornamental plants.
XX		
SQ		Sequence 385 AA;
	Query Match	92.7%; Score 1928; DB 18; Length 385;
	Best Local Similarity	100.0%; Pred. No. 2.2e-143;
	Matches 372; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
0Y	1 MSLNSTSGAGASTMOISIGAGSNNGLGTSRONAGLGGNSALGLGGGQNDPTVNOALGL	60
DB	1 MSLNSTSGAGASTMOISIGAGSNNGLGTSRONAGLGGNSALGLGGGQNDPTVNOALGL	60
0Y	61 TGNMNMMSMGGGGGLMGGLGGGLGNGLGGSGGLGEGLSNALNDMLGGSLNTLGSKGGNN	120
DB	61 TGNMNMMSMGGGGGLMGGLGGGLGNGLGGSGGLGEGLSNALNDMLGGSLNTLGSKGGNN	120
0Y	121 TTSTNTNSPLDQALGINTSTSONDSTSGDSTSDSSDPQQLLKMFSEIMQSLFGGQDGT	180
DB	121 TTSTNTNSPLDQALGINTSTSONDSTSGDSTSDSSDPQQLLKMFSEIMQSLFGGQDGT	180
0Y	181 QGSSSGGQPTEGEDONAYKKVGYDALSLGIMNGLSQQLLNGGLGGGOGGNACTGDDSSL	240
DB	181 QGSSSGGQPTEGEDONAYKKVGYDALSLGIMNGLSQQLLNGGLGGGOGGNACTGDDSSL	240
0Y	241 GKGGLQNTSGPYDVOQLNNAVGTGIGMRAGIQALNDIGTHRSSSTRSPFNKGDRAAKEI	300
DB	241 GKGGLQNTSGPYDVOQLNNAVGTGIGMRAGIQALNDIGTHRSSSTRSPFNKGDRAAKEI	300
0Y	301 GGFPMQYFVVRGKPOYOKGPGQEVKTTDDKSNKAKLSKPDODGMPASMEQFNKAGMTR	360
DB	301 GGFPMQYFVVRGKPOYOKGPGQEVKTTDDKSNKAKLSKPDODGMPASMEQFNKAGMTR	360

OY		361 PMAGDTGNGNLQ 372	
DB		361 PMAGDTGNGNLQ 372	
 RESULT 12			
ID	AAR45751		
XX	AAR45751 standard; Protein; 385 AA.		
AC	AAR45751;		
XX			
DE	11-JUL-1994 (first entry).		
XX			
DE	Ervania amylovora harpin.		
KW	Harpin: hypersensitive response elicitor; HR-elicitor; fire blight;		
RW	Rosaceae: apple; pear; phytopathogenic bacterium; defence reaction;		
KM	hypersensitive reaction and pathogenicity; hrpn; gene cluster; ds.		
XX			
OS	Ervania amylovora.		
XX			
FH	Key	Location/Qualifiers	
FT	Misc-difference 372	/note= "corresponds to CAG codon"	
FT	Misc-difference 373	/note= "corresponds to CAC codon"	
FT			
PN	WO9401546-A.		
XX			
PD	20-JAN-1994.		
PF	30-JUN-1993; 93WO-US06243.		
XX			
PR	01-JUL-1992; 92US-0907935.		
XX			
PA	(CORR) CORNELL RES FOUND INC.		
PI	Bauer DM, Beer SV, Collmer A, He S, Lady R, Wei Z;		
XX			
DR	WP1; 1994-035054/04.		
XX	N-PSDB; AAO55751.		
PT	Hypersensitive response elicitor protein derived from Erwinia		
PT	amylovora - and DNA encoding it, useful for developing harpin		
PT	inhibitors to prevent e.g. fire blight of fruit		
PS	Claim 6; Page 27-28; 47pp; English.		
XX			
CC	The hrpn gene was isolated from E.amylovora using a 48-fold		
CC	degenerate oligonucleotide probe corresponding to amino acids 9-15		
CC	at the N-terminus of harpin. The 44kd protein encoded by the hrpn		
CC	gene is a hypersensitive response elicitor protein. The harpin is		
CC	thought to be an archetype for HR elicitors from phytopathogenic		
CC	bacteria.		
XX			
SQ	Sequence 385 AA:		
Query Match	92.0%; Score 1913; DB 15; Length 385;		
Best Local Similarity	99.5%; Pred. No. 3.3e-14;		
Matches 369; Conservative	1; Mismatches 1; Indels 0; Gaps 0;		
OY			
DB			
61	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
121	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
OY			
DB			
61	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
121	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
OY			
DB			
61	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
121	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
OY			
DB			
61	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
121	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
OY			
DB			
61	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
121	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
OY			
DB			
61	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
121	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
OY			
DB			
61	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
121	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
OY			
DB			
61	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
121	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
OY			
DB			
61	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
121	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
OY			
DB			
61	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
121	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
OY			
DB			
61	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
121	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
OY			
DB			
61	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
121	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
OY			
DB			
61	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
121	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
OY			
DB			
61	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
121	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
OY			
DB			
61	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	
121	TGMMAMMSMGGGGLMGGLGGGGLGSGGLGEGLSNLNDMLGSLNTMGSKGGNN	120	


```

OY 181 QGSSSGGKQPTGEGONAYKKVTYDALSGLMNGLSQLLGNGLAGGGGCGAGTGLDGSSTL 240
    |||||
DB 181 QGSSSGGKQPTGEGONAYKKVTYDALSGLMNGLSQLLGNGLAGGGGCGAGTGLDGSSTL 240
OY 241 GKGGLQNTSGPVYDQQLGNVGTGIGMAAGIOLANDIGTHRSSSTRSFYNNKDDRAMAKKI 300
    |||||
DB 241 GKGGLQNTSGPVYDQQLGNVGTGIGMAAGIOLANDIGTHRSSSTRSFYNNKDDRAMAKKI 300
OY 301 GGFMDQYEFVEFGKPYOKGPGQEVETDDKSNPAKLSKPPDDGCTPASMEOFNKAKGMIR 360
    |||||
DB 301 GGFMDQYEFVEFGKPYOKGPGQEVETDDKSNPAKLSKPPDDGCTPASMEOFNKAKGMIR 360
OY 361 PMAGDTGNGNL 371
    |||||
DB 361 PMAGDTGNGNL 371
    |||||

RESULT 13
ID AAM06597 standard; Protein; 338 AA.
AC AAM06597;
DE 30-MAR-1997 (first entry)
XX Hypersensitive response elicitor protein.
DE Hypersensitive response elicitor protein.
KW Hypersensitive response; elicitor; Erwinia chrysanthemi; plant;
KW disease-resistance; Escherichia coli; infiltration; virus;
KW bacterium; fungus; pathogen; biological control agent.
XX Erwinia chrysanthemi.
XX OS MO9639802-A1.
XX FN 19-DEC-1996.
XX PD 05-JUN-1996; 96MO-US08819.
XX PE 07-JUN-1995; 95US-0475775.
XX PR (CORR ) CORNELL RES FOUND INC.
XX PA Beer SV, Wei Z;
XX PI MPI; 1997-051614/05.
XX DR N-PSDB; AAT49313.
XX DR Imparting pathogen resistance to plants - with hypersensitive
XX PT response elicitor polypeptide or protein
XX PS Claim 4; Page 44; 69pp; English.
XX CC This sequence represents a hypersensitive response elicitor from
XX CC Erwinia chrysanthemi, with a mol.wt. of 34,000. The elicitor is
XX CC thermostable, has a glycine content of over 16%, and has no cysteine.
XX CC The elicitor may be used in a new method for imparting pathogen
XX CC resistance to plants, by application of the elicitor in a non-
XX CC infectious form to plant cells, by spraying, injection, leaf
XX CC abrasion, or plant infection with recombinant bacteria (non-
XX CC infectious to the host plant, e.g. Escherichia coli) expressing the
XX CC elicitor as a biological control agent, to allow recombinant
XX CC protein infiltration into the plant. The method confers virus,
XX CC bacterium or fungus disease-resistance on crops and ornamental
XX CC plants.
XX PS Sequence 338 AA;
XX SQ Query Match 34.6%; Score 718.5; DB 18; Length 338;
    Best Local Similarity 42.8%; Pred. No. 18e-48;
    Matches 173; Conservative 41; Mismatches 111; Indels 79; Gaps 11;
OY 13 MOISI-GGAGGNGGLGTSRONAGLGG-NSA---LGIAGGQNDIVVNLGILLTGMNMM 67

```

```

DB 1 MOTTIAHIGGDLGVSGLAQ--GLKGLNSAASLSSVYDKLSSTIDKLTSLTSM---- 55
OY 68 SMAGGGGGLGGLGGLGKNGLGGSGGLGEGLSNALNDMLGSLINTLGSGGNNTTSTTNS 127
    |||||
DB 56 -----FGALAGGLGAS-SKGLMSNQLGOSFGN-----GAOGASNLSPK- 96
OY 128 PLDQALGINSTSGNDSTSGTSTSDSPMQQLMFPEIMQSLFG-----DGCG 179
    |||||
DB 97 -----SGDALS-----MFDAIDDLGHDPTVKLTNQSQ 128
OY 180 TQSSSGGKQPTGEGONAYKKVTYDALSGLMNGLSQLLGNGLAGGGGCGAGTGLDGSSTL 239
    |||||
DB 129 LANSMLNASQMTGNNNAFGSVNNALSLILGGLGQSM-----SGFSQS 174
OY 240 LGKGLQNTSGPVYDQQLGNVGTGIGMAAGIOLANDIGTHRSSSTRSFYNNKDDRAMAKKI 299
    |||||
DB 175 LGAGGLQNTSGPVYDQQLGNVGTGIGMAAGIOLANDIGTHRSSSTRSFYNNKDDRAMAKKI 234
OY 300 IGFMDQYEFVEFGKPYOKGPGQEVETDDKSNPAKLSKPPDDGCTPASMEOFNKAKGMIR 359
    |||||
DB 235 IGFMDQYEFVEFGKPYOKGPGQEVETDDKSNPAKLSKPPDDGCTPASMEOFNKAKGMIR 294
OY 360 RPMAGDTGNGNLQARGAGSSILGIDAMAGDAINNNAKLGAA 403
    |||||
DB 295 SAVAGDTGNTNMLRGAGGSLGIDAAVYGDRIANNSLGKLANA 338

RESULT 14
ID AAM87638 standard; Protein; 338 AA.
AC AAM87638;
DE 09-MAR-1999 (first entry)
XX A hypersensitive response elicitor protein.
DE A hypersensitive response elicitor protein.
KW Hypersensitive response elicitor protein; hairpin protein;
KW disease resistance; seed quality; insect control; corn borer;
KW Lepidoptera larvae; transgenic plant.
XX Erwinia chrysanthemi.
XX OS MO9854214-A2.
XX FN 03-DEC-1998.
XX PD 28-MAY-1998; 98MO-US10874.
XX PE 30-MAY-1997; 97US-0048109.
XX PR (CORR ) CORNELL RES FOUND INC.
XX PA (EDEN-) EDEN BIOSCIENCE CORP.
XX XX Beer SV, Laby RJ, Wei Z;
XX PI MPI; 1999-070210/06.
XX DR N-PSDB; AAV83988.
XX DR New fragments of an Erwinia hypersensitive response elicitor protein
XX PT and related DNA - used to impart disease resistance to plants, to
XX PT increase their growth and to control insects
XX PS Disclosure; Page 7-8; 94pp; English.
XX CC The present sequence represents a hypersensitive response elicitor
XX CC protein (also called hairpin protein) that is able to elicit a
XX CC hypersensitive response in plants. The specification also describes
XX CC hypersensitive response elicitors from other pathogenic organisms.
XX CC The protein, in non-infectious form, is applied to plants to impart
XX CC disease resistance (to a wide range of viral, bacterial and fungal
XX CC pathogens), to improve growth (yield, quantity and quality of seeds,

```

CC to provide earlier germination etc.) and to control insects (e.g. corn
CC borers, Lepidoptera larvae etc.) The same results are provided by
CC transgenic plants expressing the protein.

XX Sequence 338 AA:

Query Match 34.6%; Score 718.5; DB 20; Length 338;
Best Local Similarity 42.8%; Pred. No. 1.8e-48;
Matches 173; Conservative 41; Mismatches 111; Indels 79; Gaps 11;

QY 13 MQIST-GGAGGNNGLGTSRQNLGG-NSA---LGIGGNONDYVNOGLGLTGMNMM 67
DB 1 MQITRAHIGDGLGVSGIGAQ--GLKGLNSAASSLGSSVDKLSSTIDKLSALTSMV--- 55
QY 68 SMMGGGGLMGGLGGGGLGNGLGSSGGLBELSNALNDMLGSLNTLGSKGGNNTTSTNS 127
DB 56 -----FGALAGLGLAS-SKGLGMSNQLGQSFGN-----GAQASNLISVPK- 96
QY 128 PLDOALGINSTSONDSTSGTSDSPDPMQLLKMFEIMQSLFG-----DGQDG 179
DB 97 -----SGDALS-----KMFDAADDLGHDTVTTLTNQSNQ 128
QY 180 TQSSSSGGKOPTGEONAYKKVTDALSGIMNGLSQLLGNGLGGGQGGNAGTGLDSS 239
DB 129 LANSMLNASQMTQGNNAFSGSVNNAISLILGNGLGSM-----SGFSQPS 174
QY 240 LGKGLONTSGPYDYOOLGNAVGTGIGMKAGIQLANDIGHRHSSTRSFYKGDRAK 299
DB 175 LGAGGLDGLSGAGAFNOLGNALIGVGNALSLSVSTHVDGNHNFYDKEDRGMAKE 234
QY 300 IGFMDQYPEVFVKPOYOKPGQEVKTDKSNKALSKPDDDCMTSPASMEQFNKAKMIR 359
DB 235 IGFMDQYPEIFKPEYOKDGWSSPKTDKSNKALSKPDDDCMTGASMDKFNQAMGIR 294
QY 360 RPAAGTGNKNTLQARAGSSSLGIDMAMAGDAINNNAKGLGAA 403
DB 295 SAVAAGDTGNTNMLRGAGASLIGIDAAYVGDKIANNISLIGLANA 338

RESULT 15

AAW82407
ID AAW82407 standard; Protein; 338 AA.

AC AAW82407;

DT 23-FEB-1999 (first entry)

DE E. chrysanthemi hrpN-Ech protein.

KM Hypersensitive response elicitor protein; hrpN-Ech; pathogen resistance;
KM plant; transformation; pathogen-inducible promoter.

OS Erwinia chrysanthemi.

PN US5850015-A.

PD 15-DEC-1998.

PF 07-JUN-1995; 95US-0484358.

PR 07-JUN-1995; 95US-0484358.

PA (CORR) CORNELL RES FOUND INC.

PI Bauer D, Collmer A;

DR WPI; 1999-069852/06.

DR N-PSDB; AAV73494, AAV73507.

PT DNA encoding Erwinia chrysanthemi hypersensitive response protein
hrpN - useful for imparting pathogen resistance to plants

PS Claim 2; Column 29-30; 27pp; English.

XX This sequence represents a novel Erwinia chrysanthemi protein, hrpN-Ech,
CC that elicits a hypersensitive response in plants. The encoding DNA can be
CC used for imparting pathogen resistance to plants, by transforming a
CC plant with a vector containing the DNA and a pathogen-inducible promoter.

XX Sequence 338 AA:

Query Match 34.6%; Score 718.5; DB 20; Length 338;
Best Local Similarity 42.8%; Pred. No. 1.8e-48;
Matches 173; Conservative 41; Mismatches 111; Indels 79; Gaps 11;

QY 13 MQIST-GGAGGNNGLGTSRQNLGG-NSA---LGIGGNONDYVNOGLGLTGMNMM 67
DB 1 MQITRAHIGDGLGVSGIGAQ--GLKGLNSAASSLGSSVDKLSSTIDKLSALTSMV--- 55
QY 68 SMMGGGGLMGGLGGGGLGNGLGSSGGLBELSNALNDMLGSLNTLGSKGGNNTTSTNS 127
DB 56 -----FGALAGLGLAS-SKGLGMSNQLGQSFGN-----GAQASNLISVPK- 96
QY 128 PLDOALGINSTSONDSTSGTSDSPDPMQLLKMFEIMQSLFG-----DGQDG 179
DB 97 -----SGDALS-----KMFDAADDLGHDTVTTLTNQSNQ 128
QY 180 TQSSSSGGKOPTGEONAYKKVTDALSGIMNGLSQLLGNGLGGGQGGNAGTGLDSS 239
DB 129 LANSMLNASQMTQGNNAFSGSVNNAISLILGNGLGSM-----SGFSQPS 174
QY 240 LGKGLONTSGPYDYOOLGNAVGTGIGMKAGIQLANDIGHRHSSTRSFYKGDRAK 299
DB 175 LGAGGLDGLSGAGAFNOLGNALIGVGNALSLSVSTHVDGNHNFYDKEDRGMAKE 234
QY 300 IGFMDQYPEVFVKPOYOKPGQEVKTDKSNKALSKPDDDCMTSPASMEQFNKAKMIR 359
DB 235 IGFMDQYPEIFKPEYOKDGWSSPKTDKSNKALSKPDDDCMTGASMDKFNQAMGIR 294
QY 360 RPAAGTGNKNTLQARAGSSSLGIDMAMAGDAINNNAKGLGAA 403
DB 295 SAVAAGDTGNTNMLRGAGASLIGIDAAYVGDKIANNISLIGLANA 338

Search completed: June 12, 2003, 08:49:12
Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2003, 08:48:31 ; Search time 14 Seconds
(without alignments)
846.960 Million cell updates/sec

Title: US-09-412-100-23

Perfect score: 2079 1 MSMTSGIGASTMQISIGCA.....DAMAGDAINNALGKIGAA 403

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: Issued_Patents_AA:*
2: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2079	100.0	403	2	US-08-724A-2
2	2079	100.0	403	2	US-09-030-270A-3
3	2079	100.0	403	4	US-08-851-376A-2
4	2079	100.0	403	4	US-08-984-207-3
5	2079	100.0	403	4	US-09-013-587-3
6	1928	92.7	385	1	US-08-891-254-3
7	1928	92.7	385	2	US-08-819-539-3
8	1928	92.7	385	5	PCT-US96-08819-3
9	1913	92.0	385	5	PCT-US93-06243-2
10	718.5	34.6	338	1	US-08-891-254-1
11	718.5	34.6	338	2	US-08-484-358-2
12	718.5	34.6	338	2	US-08-819-539-1
13	718.5	34.6	338	2	US-09-030-270A-1
14	718.5	34.6	338	3	US-09-118-959-2
15	718.5	34.6	338	4	US-08-984-207-1
16	718.5	34.6	338	4	US-09-013-587-1
17	718.5	34.6	338	4	PCT-US96-08819-1
18	211.5	10.2	318	4	US-09-060-756-727
19	211.5	10.2	651	4	US-08-556-978B-19
20	211.5	10.2	651	4	US-09-247-806-1
21	211.5	10.2	718	1	US-08-425-069-2
22	211.5	10.2	718	2	US-08-317-844B-2
23	211.5	10.2	747	3	US-09-034-177-3
24	206	9.9	604	4	US-08-556-978B-63
25	199.5	9.6	334	4	US-09-060-756-728
26	198.5	9.5	738	3	US-08-864-038A-3
27	198	9.5	606	4	US-09-247-806-6

28	197.5	9.5	344	1	US-08-891-254-7	Sequence 7, Appl
29	197.5	9.5	344	2	US-08-819-539-7	Sequence 7, Appl
30	197.5	9.5	344	2	US-09-030-270A-7	Sequence 7, Appl
31	197.5	9.5	344	4	US-08-984-207-7	Sequence 7, Appl
32	197.5	9.5	344	4	US-09-013-587-7	Sequence 7, Appl
33	197.5	9.5	344	5	PCT-US96-08819-7	Sequence 7, Appl
34	197	9.5	606	4	US-08-556-978B-23	Sequence 23, Appl
35	197	9.5	606	4	US-09-247-806-8	Sequence 8, Appl
36	194	9.3	1160	3	US-08-808-598A-24	Sequence 24, Appl
37	192.5	9.3	606	4	US-08-556-978B-21	Sequence 21, Appl
38	192.5	9.3	606	4	US-09-247-806-4	Sequence 4, Appl
39	190	9.1	641	1	US-09-249-585A-3	Sequence 3, Appl
40	186	8.9	749	1	US-08-317-522A-2	Sequence 2, Appl
41	186	8.9	749	1	US-08-439-818A-2	Sequence 2, Appl
42	186	8.9	749	2	US-08-751-965-2	Sequence 2, Appl
43	186	8.9	749	2	US-08-738-975-2	Sequence 2, Appl
44	186	8.9	749	2	US-08-728-626-2	Sequence 2, Appl
45	186	8.9	749	3	US-08-808-599A-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-200-724A-2
; Sequence 2, Application US/08200724A
; Patent No. 5849868
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Bauer, David W.
; APPLICANT: Beer, Steven V.
; APPLICANT: Collier, Alan
; APPLICANT: He, Sheng-Yang
; APPLICANT: Laby, Ron J.
; TITLE OF INVENTION: ELICITOR OF THE HYPERSENSITIVE RESPONSE
; NUMBER OF SEQUENCES: 5
; TITLE OF INVENTION: IN PLANTS
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentia Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/200,724A
; FILING DATE: 23-FEB-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/10030
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1304
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 403 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-200-724A-2
; Query Match 100.0%; Score 2079; DB 2; Length 403;
; Best Local Similarity 100.0%; Pred. No. 2.4e-174;
; Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 403 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-851-376A-2

Query Match 100.0%; Score 2079; DB 4; Length 403;
Best Local Similarity 100.0%; Pred. No. 2.4e-174;
Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLSLTSGLASTMSTISIGAGGNNGLGTSRONAGLGSNSALGLGGNDTVNOLAGLL 60
DB 1 MSLSLTSGLASTMSTISIGAGGNNGLGTSRONAGLGSNSALGLGGNDTVNOLAGLL 60
QY 61 TGMAMMSMMGGGGLMGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGL 120
DB 61 TGMAMMSMMGGGGLMGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGL 120
QY 121 TTSTNSPLDQALINSTSQNDSTSGTSDSDSPMOQLKMFSEIMOSLFGDGQDCT 180
DB 121 TTSTNSPLDQALINSTSQNDSTSGTSDSDSPMOQLKMFSEIMOSLFGDGQDCT 180
QY 181 QGSSSGGKOPTGEGONAKKVTDLALSLMGNGLSQLLGNGLGGGCGGNAGTGLDGSSL 240
DB 181 QGSSSGGKOPTGEGONAKKVTDLALSLMGNGLSQLLGNGLGGGCGGNAGTGLDGSSL 240
QY 241 GGGKGLNLSGPDYQOLGNAVGTGIGMKAGIQALNDIGTHRSSSTRSFVNGDRAMAEI 300
DB 241 GGGKGLNLSGPDYQOLGNAVGTGIGMKAGIQALNDIGTHRSSSTRSFVNGDRAMAEI 300
QY 301 GGFMDQYPEVFGKPOYQKPGQEVKTDKSNAAKALSKPDDDMTPASMEQFNKAKGMIR 360
DB 301 GGFMDQYPEVFGKPOYQKPGQEVKTDKSNAAKALSKPDDDMTPASMEQFNKAKGMIR 360
QY 361 PMAGDTGNGNLQARGAGSSSLGIDAMMAGDAIINNALGKLGAA 403
DB 361 PMAGDTGNGNLQARGAGSSSLGIDAMMAGDAIINNALGKLGAA 403

RESULT 4

US-08-984-207-3
Sequence 3, Application US/08984207
Patent No. 6235974
GENERAL INFORMATION:

APPLICANT: Olu, Deven
APPLICANT: Wei, Zhong-Min
APPLICANT: Beer, Steven V.
TITLE OF INVENTION: HYPERSENSITIVE RESPONSE INDUCED
TITLE OF INVENTION: RESISTANCE IN PLANTS BY SEED TREATMENT
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:

ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: P.O. Box 1051, Clinton Square
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/984,207
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/033,230
FILING DATE: 05-DEC-1996

ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/1201
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 403 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-984-207-3

Query Match 100.0%; Score 2079; DB 4; Length 403;
Best Local Similarity 100.0%; Pred. No. 2.4e-174;
Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLSLTSGLASTMSTISIGAGGNNGLGTSRONAGLGSNSALGLGGNDTVNOLAGLL 60
DB 1 MSLSLTSGLASTMSTISIGAGGNNGLGTSRONAGLGSNSALGLGGNDTVNOLAGLL 60
QY 61 TGMAMMSMMGGGGLMGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGL 120
DB 61 TGMAMMSMMGGGGLMGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGL 120
QY 121 TTSTNSPLDQALINSTSQNDSTSGTSDSDSPMOQLKMFSEIMOSLFGDGQDCT 180
DB 121 TTSTNSPLDQALINSTSQNDSTSGTSDSDSPMOQLKMFSEIMOSLFGDGQDCT 180
QY 181 QGSSSGGKOPTGEGONAKKVTDLALSLMGNGLSQLLGNGLGGGCGGNAGTGLDGSSL 240
DB 181 QGSSSGGKOPTGEGONAKKVTDLALSLMGNGLSQLLGNGLGGGCGGNAGTGLDGSSL 240
QY 241 GGGKGLNLSGPDYQOLGNAVGTGIGMKAGIQALNDIGTHRSSSTRSFVNGDRAMAEI 300
DB 241 GGGKGLNLSGPDYQOLGNAVGTGIGMKAGIQALNDIGTHRSSSTRSFVNGDRAMAEI 300
QY 301 GGFMDQYPEVFGKPOYQKPGQEVKTDKSNAAKALSKPDDDMTPASMEQFNKAKGMIR 360
DB 301 GGFMDQYPEVFGKPOYQKPGQEVKTDKSNAAKALSKPDDDMTPASMEQFNKAKGMIR 360
QY 361 PMAGDTGNGNLQARGAGSSSLGIDAMMAGDAIINNALGKLGAA 403
DB 361 PMAGDTGNGNLQARGAGSSSLGIDAMMAGDAIINNALGKLGAA 403

RESULT 5

US-09-013-587-3
Sequence 3, Application US/09013587
Patent No. 6277814
GENERAL INFORMATION:

APPLICANT: Olu, Deven
APPLICANT: Wei, Zhong-Min
APPLICANT: Beer, Steven V.
TITLE OF INVENTION: ENHANCEMENT OF GROWTH IN PLANTS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:

ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/013,587

FILED DATE: 27-JAN-1997
CLASSIFICATION: US 60/036,048
PRIOR APPLICATION DATA: 27-JAN-1997
FILING DATE: 27-JAN-1997
ATTORNEY/AGENT INFORMATION: NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/1501
TELECOMMUNICATION INFORMATION: TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 403 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-013-587-3

Query Match 100.0%; Score 2079; DB 4; Length 403;
Best Local Similarity 100.0%; Pred. No. 2,4e-174;
Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLTSGIGASTMQLISIGAGGNNGLCTSRQNAIGLGSNLSALGCGNQNDPTVNLGIL 60
DB 1 MSLTSGIGASTMQLISIGAGGNNGLCTSRQNAIGLGSNLSALGCGNQNDPTVNLGIL 60
QY 61 TGMAMMAMMGGGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGL 120
DB 61 TGMAMMAMMGGGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGL 120
QY 121 TTSTNSPLDALGINSTSONDSTSGTDSSTSDPMQOLLKMFSEIMQSLFGGQDGT 180
DB 121 TTSTNSPLDALGINSTSONDSTSGTDSSTSDPMQOLLKMFSEIMQSLFGGQDGT 180
QY 181 OGSSSGKOPTREGONAKKGYTDALSGIMNGLSQLLGNGLGCGGCGNAGTGLDSSL 240
DB 181 OGSSSGKOPTREGONAKKGYTDALSGIMNGLSQLLGNGLGCGGCGNAGTGLDSSL 240
QY 241 GKGGLQNTSGPVDYQQLGNAGVTGIGMAGIOLANDIGTHRSSSTRFVNGDRMAKEI 300
DB 241 GKGGLQNTSGPVDYQQLGNAGVTGIGMAGIOLANDIGTHRSSSTRFVNGDRMAKEI 300
QY 301 GGFMDQYEVFGKPGYQKPGQEVKTDKSNKALSKRPDDGMPASMEQFNKAKGMIKR 360
DB 301 GGFMDQYEVFGKPGYQKPGQEVKTDKSNKALSKRPDDGMPASMEQFNKAKGMIKR 360
QY 361 PMAGDTGNGNLQARGAGSSSLGIDAMMAGDAINNNALGKLGAA 403
DB 361 PMAGDTGNGNLQARGAGSSSLGIDAMMAGDAINNNALGKLGAA 403

RESULT 6
US-08-891-254-3
Sequence 3, Application US/08891254
Patent No. 5776889
GENERAL INFORMATION:
APPLICANT: Wei, Zhong-Min
APPLICANT: Beer, Steven V.
TITLE OF INVENTION: Hypersensitive Response
TITLE OF INVENTION: Induced Resistance in Plants
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/891,254
FILING DATE: 10-JUL-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/475,775
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 14603/10050
TELECOMMUNICATION INFORMATION: TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-891-254-3

Query Match 92.7%; Score 1928; DB 1; Length 385;
Best Local Similarity 100.0%; Pred. No. 3,8e-161;
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLTSGIGASTMQLISIGAGGNNGLCTSRQNAIGLGSNLSALGCGNQNDPTVNLGIL 60
DB 1 MSLTSGIGASTMQLISIGAGGNNGLCTSRQNAIGLGSNLSALGCGNQNDPTVNLGIL 60
QY 61 TGMAMMAMMGGGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGL 120
DB 61 TGMAMMAMMGGGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGL 120
QY 121 TTSTNSPLDALGINSTSONDSTSGTDSSTSDPMQOLLKMFSEIMQSLFGGQDGT 180
DB 121 TTSTNSPLDALGINSTSONDSTSGTDSSTSDPMQOLLKMFSEIMQSLFGGQDGT 180
QY 181 OGSSSGKOPTREGONAKKGYTDALSGIMNGLSQLLGNGLGCGGCGNAGTGLDSSL 240
DB 181 OGSSSGKOPTREGONAKKGYTDALSGIMNGLSQLLGNGLGCGGCGNAGTGLDSSL 240
QY 241 GKGGLQNTSGPVDYQQLGNAGVTGIGMAGIOLANDIGTHRSSSTRFVNGDRMAKEI 300
DB 241 GKGGLQNTSGPVDYQQLGNAGVTGIGMAGIOLANDIGTHRSSSTRFVNGDRMAKEI 300
QY 301 GGFMDQYEVFGKPGYQKPGQEVKTDKSNKALSKRPDDGMPASMEQFNKAKGMIKR 360
DB 301 GGFMDQYEVFGKPGYQKPGQEVKTDKSNKALSKRPDDGMPASMEQFNKAKGMIKR 360
QY 361 PMAGDTGNGNLQ 372
DB 361 PMAGDTGNGNLQ 372

RESULT 7
US-08-819-539-3
Sequence 3, Application US/08819539
Patent No. 5859324
GENERAL INFORMATION:
APPLICANT: Wei, Zhong-Min
APPLICANT: Beer, Steven V.
TITLE OF INVENTION: Hypersensitive Response
TITLE OF INVENTION: Induced Resistance in Plants
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester

STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 05/08/919,539
FILING DATE: 17-MAR-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/475,775
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 14603/10050
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
type: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-619-539-3

ADDRESSEE: Nixon, Hairgrave, Devans & Doyle Ltd
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCV/US96/08819
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/475,775
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/10051
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: Linear
MOLECULE TYPE: protein
PCT-US96-08819-3


```

; TITLE OF INVENTION: Elicitor of the Hypersensitive Response in Plants
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06243
; FILING DATE: 19930630
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 907,935
; FILING DATE: 01-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak
; REGISTRATION NUMBER: 26,824
; REFERENCE/DOCKET NUMBER: CRF D-1172
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203)268-1951
; TELEFAX: (203)268-1951
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 385 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US93-06243-2

Query Match      92.0%; Score 1913; DB 5; Length 385;
Best Local Similarity 99.5%; Pred. No. 7.9e-160;
Matches 369; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSINTSGASTMWSISGAGGNGLLGSTRONAGLGAGNAGLGGGNDYNOAGLL 60
DB 1 MSINTSGASTMWSISGAGGNGLLGSTRONAGLGAGNAGLGGGNDYNOAGLL 60
QY 61 TGMAMMAMMAGGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGL 120
DB 61 TGMAMMAMMAGGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGL 120
QY 121 TSTSTNSPLDQALGINSTSONDSTSGTSTSDSSDPMOQLKMFSEIMOSLFGDGDGT 180
DB 121 TSTSTNSPLDQALGINSTSONDSTSGTSTSDSSDPMOQLKMFSEIMOSLFGDGDGT 180
QY 181 QCSSSSGKQPTREGEONAYKRYTDALSGLMGNGLSOLLGNGLGAGGCGGAGTGLDSSSL 240
DB 181 QCSSSSGKQPTREGEONAYKRYTDALSGLMGNGLSOLLGNGLGAGGCGGAGTGLDSSSL 240
QY 241 GKGGLGNTSGPVDYQOLGNAVGTGIGMKAGIOLANDIGTHRSSTSFVKKGRAMAKEI 300
DB 241 GKGGLGNTSGPVDYQOLGNAVGTGIGMKAGIOLANDIGTHRSSTSFVKKGRAMAKEI 300
QY 301 GGFPMOYRPFVFGKPOYQKPGQGVTKTDDSKMAKALSKPDDDGMTFPAISMEQFNRAKGMIR 360
DB 301 GGFPMOYRPFVFGKPOYQKPGQGVTKTDDSKMAKALSKPDDDGMTFPAISMEQFNRAKGMIR 360
QY 361 PMAGDTGNGNL 371
DB 361 PMAGDTGNGNL 371

RESULT 10
US-08-891-254-1
; Sequence 1, Application US/08891254

```

```

; Patent No. 5776889
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Beer, Steven V.
; TITLE OF INVENTION: Hypersensitive Response
; TITLE OF INVENTION: Induced Resistance in Plants
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,254
; FILING DATE: 10-JUL-1997
; CLASSIFICATION: 514
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/475,775
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 14603/10050
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1304
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 338 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-891-254-1

Query Match      34.6%; Score 718.5; DB 1; Length 338;
Best Local Similarity 42.8%; Pred. No. 2.8e-55;
Matches 173; Conservative 41; Mismatches 11; Indels 79; Gaps 11;

QY 13 MOISI-GGAGNGNGLGSTRONAGLG--LGLGNGNDYNOAGLLTGMAMMM 67
DB 1 MOITTAHIGGDLVSGLGMO--GLGLNMAASLSGVKLSSTIDKTSALTSMV--- 55
QY 68 SMAGGGGLMGGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGLGGL 127
DB 56 -----FGGALAGGLGAS-SKGLMSHQLGOSFEN-----GAQASMLSVPK- 96
QY 128 PLDQALGINSTSONDSTSGTSTSDSSDPMOQLKMFSEIMOSLFG-----DQDGG 179
DB 97 -----SGGDALS-----KMFALADLDLGHVTVTLTNSNQ 128
QY 180 TQSSSSGKQPTREGEONAYKRYTDALSGLMGNGLSOLLGNGLGAGGCGGAGTGLDSS 239
DB 129 LANSMLNASQOTGNNMAFGSSGVNNAISLILGGLGOSM-----SGFSOPS 174
QY 240 LGKGLGNTSGPVDYQOLGNAVGTGIGMKAGIOLANDIGTHRSSTSFVKKGRAMAKE 299
DB 175 LGAGGLGNTSGAFAFNDLGNATGMYGQNALNLSVSTHYVGNRHVYDKEDRGMAKE 234
QY 300 IGFPMOYRPFVFGKPOYQKPGQGVTKTDDSKMAKALSKPDDDGMTFPAISMEQFNRAKGMIR 359
DB 235 IGFPMOYRPFVFGKPOYQKPGQGVTKTDDSKMAKALSKPDDDGMTGASMDKFRQAGMIR 294
QY 360 RPAAGDTGNGNLQARGAGSSSLGTDAMMGDALNNNAIGLGA 403
DB 295 SAVAGDTGNTNLNRGAGGASLIGIDAAYVGDKTANMSLGLANA 338

```

RESULT 11

US-08-484-358-2

Sequence 2, Application US/08484358

Patent No. 5850015

GENERAL INFORMATION:

APPLICANT: Bauer, David

APPLICANT: Collmer, Alan

TITLE OF INVENTION: Hypersensitive Response Elicitor

TITLE OF INVENTION: From

TITLE OF INVENTION: Erwinia Chrysanthemi

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Nixon, Hargrave, Devans & Doyle

STREET: Clinton Square

CITY: Rochester

STATE: New York

COUNTRY: U.S.A.

ZIP: 14603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/484,358

FILING DATE:

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Goldman, Michael L.

REGISTRATION NUMBER: 30,727

REFERENCE/DOCKET NUMBER: 19603/840

TELECOMMUNICATION INFORMATION:

TELEPHONE: 716-263-1304

TELEFAX: 716-263-1600

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 338 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-484-358-2

Query Match 34.6%; Score 718.5; DB 2; Length 338;
Best Local Similarity 42.8%; Pred. No. 2.8e-55;
Matches 173; Conservative 41; Mismatches 111; Indels 79; Gaps 11;

QY 13 MOIST-GGAGGNNGLGTSRONAGLGG-NSA---LGIGGNNNDVYNQAGLITGMMMM 67
DB 1 MOITIKAHIGDGLGSLGAQ--GLKGLNSAASSLSSVDKLSITDKTSLTSM--- 55
QY 68 SMWGCGGLMGGLGGLGSLGSGGLGSLNGLNGLGSLNGLGSGKGNNTSTNS 127
DB 56 -----FGALAGGLGAS-SKGLGKSNQGLQSGN-----GAGGASNLISVPR- 96
QY 128 PLDQALGINSTSUNDSTSGDSTSDSDPMOQLKMFSEINQSLFG-----DGQDG 179
DB 97 -----SGDALS-----KMFKALDDLGHPTVTKLNQSNQ 128
QY 180 TCGSSSGGKQPTBEGDQNAKKGVTDALSGLMGNGLSQLLGLNGSLGGGCGGAGNAGTGLDGS 239
DB 129 LANSMLNASQMTQGNNAFSGSVNNALSIILNGLGQSM-----SGFSOPS 174
QY 240 LGGKGLONTSGPVDYQOLGNAVGTGIGMAGAGIOLNDICTHRRSSTRSPVNGGDRAMAKE 299
DB 175 LGAGGLGSLGAGAFNQLGNAIGMGVQNAALSLALSNVSTHYDGNRRHFDKEDRGMAKE 234
QY 300 IGFQFDQIPEYVFGKFOYQKGPQGVKTTDKRSNAKLSKPDGKTPASMEQFNKAKGMIT 359
DB 235 IGFQFDQIPEYVFGKFOYQKGPQGVKTTDKRSNAKLSKPDGKTPASMEQFNKAKGMIT 294

QY 360 RMAPDGTGNGNLQARGAGSSSLGIDAMAGALINMMAIGKLGAA 403
DB 295 SAVAGDGTGNTNLNLKAGAGASLGIDAAYVGDKIANNISGLKLANA 338

RESULT 12

US-08-819-539-1

Sequence 1, Application US/08819539

Patent No. 5859324

GENERAL INFORMATION:

APPLICANT: Wei, Zhong-Min

APPLICANT: Beer, Steven V.

TITLE OF INVENTION: Hypersensitive Response

TITLE OF INVENTION: Induced Resistance in Plants

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Nixon, Hargrave, Devans & Doyle

STREET: Clinton Square, P.O. Box 1051

CITY: Rochester

STATE: New York

COUNTRY: U.S.A.

ZIP: 14603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/819,539

FILING DATE: 17-MAR-1997

CLASSIFICATION: 800

PRIOR APPLICATION NUMBER: 08/475,775

ATTORNEY/AGENT INFORMATION:

NAME: Goldman, Michael L.

REGISTRATION NUMBER: 30,727

REFERENCE/DOCKET NUMBER: 14603/10050

TELECOMMUNICATION INFORMATION:

TELEPHONE: (716) 263-1304

TELEFAX: (716) 263-1600

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 338 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-819-539-1

Query Match 34.6%; Score 718.5; DB 2; Length 338;
Best Local Similarity 42.8%; Pred. No. 2.8e-55;
Matches 173; Conservative 41; Mismatches 111; Indels 79; Gaps 11;

QY 13 MOIST-GGAGGNNGLGTSRONAGLGG-NSA---LGIGGNNNDVYNQAGLITGMMMM 67
DB 1 MOITIKAHIGDGLGSLGAQ--GLKGLNSAASSLSSVDKLSITDKTSLTSM--- 55
QY 68 SMWGCGGLMGGLGGLGSLGSGGLGSLNGLNGLGSLNGLGSGKGNNTSTNS 127
DB 56 -----FGALAGGLGAS-SKGLGKSNQGLQSGN-----GAGGASNLISVPR- 96
QY 128 PLDQALGINSTSUNDSTSGDSTSDSDPMOQLKMFSEINQSLFG-----DGQDG 179
DB 97 -----SGDALS-----KMFKALDDLGHPTVTKLNQSNQ 128
QY 180 TCGSSSGGKQPTBEGDQNAKKGVTDALSGLMGNGLSQLLGLNGSLGGGCGGAGNAGTGLDGS 239
DB 129 LANSMLNASQMTQGNNAFSGSVNNALSIILNGLGQSM-----SGFSOPS 174
QY 240 LGGKGLONTSGPVDYQOLGNAVGTGIGMAGAGIOLNDICTHRRSSTRSPVNGGDRAMAKE 299
DB 175 LGAGGLGSLGAGAFNQLGNAIGMGVQNAALSLALSNVSTHYDGNRRHFDKEDRGMAKE 234

QY 300 IGFMDQPEYFGRKPOYKPGSGVEKTDKRSMAKLSRPDDGKTPASMEQFNKAKMIK 359
DB 235 IGFMDQPEYFGRKPOYKPGSGVEKTDKRSMAKLSRPDDGKTPASMEQFNKAKMIK 294
QY 360 RPAAGDTGNGNLTQARGAGSSSLGIDAMAGDAINNMALGKIGAA 403
DB 295 SAVAGDTGNTNLTNRGAGGASLGIDAAYVGDKNKINMSLGKLANA 338

RESULT 13

US-09-030-270A-1

Sequence 1, Application US/09030270A

Patent No. 5977060

GENERAL INFORMATION:

APPLICANT: Zitter, Thomas A.

APPLICANT: Wei, Zhong-Min

TITLE OF INVENTION: INSECT CONTROL WITH A

TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP

STREET: P.O. Box 1051, Clinton Square

CITY: Rochester

STATE: New York

COUNTRY: U.S.A.

ZIP: 14603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/030,270A

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/039,226

FILING DATE: 28-FEB-1997

ATTORNEY/AGENT INFORMATION:

NAME: Goldman, Michael L.

REGISTRATION NUMBER: 30,727

REFERENCE/DOCKET NUMBER: 19603/1521

TELECOMMUNICATION INFORMATION:

TELEPHONE: (716) 263-1304

TELEFAX: (716) 263-1600

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 338 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-030-270A-1

Query Match

Best Local Similarity 34.6%; Score 718.5; DB 2; Length 338;

Matches 173; Conservative 41; Mismatches 111; Indels 79; Gaps 11;

QY 13 MOIST-GGAGGNNILGTSRONAGIG-NSA---LGAGGQNDVTNOLAGLTGMMMM 67
DB 1 MOITIKAHIGDGLVSGIAQ--GLKGLNSAASSILGSSVDTLSITIDLTSLTSM--- 55
QY 68 SMAGGGGLMGGGGLGAGGGLGAGGSGGEGISNLMNDMLGSLMTLTSKSGKNNTTSTNS 127
DB 56 -----FEGALAAQGLGAS-SKGLGMSNOLGQSGFN-----GAQGASNLLSVPK- 96
QY 128 PLDQALGINSTQNDSTSGTSTSDSDPMQQLKMFSEIMQSLFG-----DQDQ 179
DB 97 -----SGGDALS-----KMFDRKALDLDLGHDTVTKLTNOSNQ 128
QY 180 TGGSSSGKQPTFEEQNAKKGVTDALSGLMGNGLSQLLGNGGLGAGGAGNAGTCLDSS 239
DB 180 TGGSSSGKQPTFEEQNAKKGVTDALSGLMGNGLSQLLGNGGLGAGGAGNAGTCLDSS 239

DB 129 LANSNLNSOMTQGMNAFSGCVNNALSSILGNGIGQSN-----SGFSQPS 174
QY 240 LGKGLQNLGSPVDYQOGLNNAVGTGIGMAGIOLNDICTHRSSRSPVNGDPAKKE 299
DB 175 LGAGLQGLGSGAGVNOGLNMGVGNALSLSNVSTHVDGNNRHFVDEKGMKE 234
QY 300 IGFMDQPEYFGRKPOYKPGSGVEKTDKRSMAKLSRPDDGKTPASMEQFNKAKMIK 359
DB 295 SAVAGDTGNTNLTNRGAGGASLGIDAAYVGDKNKINMSLGKLANA 338
QY 360 RPAAGDTGNGNLTQARGAGSSSLGIDAMAGDAINNMALGKIGAA 403
DB 295 SAVAGDTGNTNLTNRGAGGASLGIDAAYVGDKNKINMSLGKLANA 338

RESULT 14

US-09-118-959-2

Sequence 2, Application US/09118959

Patent No. 6001959

GENERAL INFORMATION:

APPLICANT: Bauer, David

APPLICANT: Collier, Alan

TITLE OF INVENTION: Hypersensitive Response Elicitor From

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Nixon, Hargrave, Devans & Doyle

STREET: Clinton Square

CITY: Rochester

STATE: New York

COUNTRY: U.S.A.

ZIP: 14603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/118,959

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Goldman, Michael L.

REGISTRATION NUMBER: 30,727

REFERENCE/DOCKET NUMBER: 19603/840

TELECOMMUNICATION INFORMATION:

TELEPHONE: 716-263-1304

TELEFAX: 716-263-1600

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 338 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-118-959-2

Query Match

Best Local Similarity 34.6%; Score 718.5; DB 3; Length 338;

Matches 173; Conservative 41; Mismatches 111; Indels 79; Gaps 11;

QY 13 MOIST-GGAGGNNILGTSRONAGIG-NSA---LGAGGQNDVTNOLAGLTGMMMM 67
DB 1 MOITIKAHIGDGLVSGIAQ--GLKGLNSAASSILGSSVDTLSITIDLTSLTSM--- 55
QY 68 SMAGGGGLMGGGGLGAGGGLGAGGSGGEGISNLMNDMLGSLMTLTSKSGKNNTTSTNS 127
DB 56 -----FEGALAAQGLGAS-SKGLGMSNOLGQSGFN-----GAQGASNLLSVPK- 96
QY 128 PLDQALGINSTQNDSTSGTSTSDSDPMQQLKMFSEIMQSLFG-----DQDQ 179
DB 97 -----SGGDALS-----KMFDRKALDLDLGHDTVTKLTNOSNQ 128
QY 180 TGGSSSGKQPTFEEQNAKKGVTDALSGLMGNGLSQLLGNGGLGAGGAGNAGTCLDSS 239
DB 180 TGGSSSGKQPTFEEQNAKKGVTDALSGLMGNGLSQLLGNGGLGAGGAGNAGTCLDSS 239

QY 180 TCGSSSGKRPKEGGGNNATKFCVTYATLSLIMGNGLSOLLGNGSLGGGGGGMNATGIDGSS 239
 Db 129 LANSMLNMSQMTQGMNNAFSGSGVNNALSSITLNGLSQSM-----SGSPSPS 174
 QY 240 LGGKGLQNLSPGVHYQQLGNNAVGTGIGMAGIGIALNDICTGHRHSYRSPFYNKGDRLMAKE 299
 Db 175 LGAGGLGTLSCGAENFQNLGNALIGMGVGNALSLNSVSTHYDGNRRHFYDKDRGMARE 234
 QY 300 IGGFNDQYPEYVGRKQYQKGPQGEVKTDDKSNAAKLSKPDDDGMPDASMEQFNKATGMK 359
 Db 235 IGGFNDQYPEYVGRKQYQKGPQGEVKTDDKSNAAKLSKPDDDGMPDASMDKFPQAGMIR 294
 QY 360 RPMAGDTGNGNLQARGAGSSGLGIDPAMAGDAILNNALRLTGA 403
 Db 295 SAVAGDTGNTNINLARGAGASLGLDPAVAGDRTANNSLRLTANA 338

RESULT 15
US-08-984-207-1

```

; Sequence 1, Application US/08984207
; Patent No. 6235974
; GENERAL INFORMATION:
;

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

ADDRESS: Nixon, Hargrave, Devans & Doyle LLP
STREET: P.O. Box 1051, Clinton Square
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603

```

1 COMPUTER READABLE FORM:
2 MEDIUM TYPE: Floppy disk
3 COMPUTER: IBM PC compatible
4 OPERATING SYSTEM: PC-DOS/MS-DOS
5 SOFTWARE: Patent In Release #1.0, Version #1.30
6 CURRENT APPLICATION DATA:
7 APPLICATION NUMBER: US/08/984,207
8 FILING DATE:

```

```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/033,230
; FILING DATE: 03/03/2000

```

FILING DATE: 05-DEC-1996
ATTORNEY/AGENT INFORMATION:

NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/1201
TELECOMMUNICATIONS INFORMATION:

TELEPHONE: (716) 263-1304
TELETYPE: (716) 263-1500

TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 1:
SEQUENCE AND ORIGIN

SEQUENCE CHARACTERISTICS:
LENGTH: 338 amino acids

```

; amino acid
; type:
; STRANDEDNESS:

```

```

;      TOPOLOGY: linear
;      MOLECULE TYPE: protein
;

```

US-08-984-207-1

Query Match	34.6%	Score 718.5;	DB 4;	Length 338;
Best Local Similarity	42.8%	Pred. No. 2.8e-55;		
Matches 173; Conservative	41;	Mismatches 111;	Indels 79;	Gaps 11;

13 MOISI-GGAGNNGILGTSRONAGLG-NSA---LGGGGNDTVNQLAGLITGMMMM 67

Db
1 MQITKAHIGDLGVSGIAQ--GLKGLNSAASSLSSVDKLSITIDKLTSALTSM--- 55

68 SMGGGGLMGGLGGGLGNGLGSSGGLGEGLSNALNDMLGSLNTLGSKGNNITTS^WTNS 127

[illegible]

Search completed: June 12, 2003, 08:50:59
Job time : 16 secs

Job time : 16 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2003, 08:49:16 ; Search time 22 seconds
(without alignments)
1891.174 Million cell updates/sec

Title: US-09-412-100-23

Perfect score: 2079

Sequence: 1 MSLSITGSLASTMQLSISGA.....DAMAGDAINNALKRLGNA 403

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PC7_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/PC705_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2079	100.0	403	9	US-10-034-158-3
2	2079	100.0	403	9	US-09-086-118-23
3	2079	100.0	403	10	US-09-835-684-3
4	2079	100.0	403	10	US-09-880-371-3
5	2079	100.0	403	10	US-09-879-248-3
6	2079	100.0	403	10	US-09-770-693-3
7	2079	100.0	403	10	US-09-766-348-3
8	718.5	34.6	338	10	US-09-086-118-21
9	718.5	34.6	338	10	US-09-835-684-1
10	718.5	34.6	338	10	US-09-880-371-1
11	718.5	34.6	338	10	US-09-879-248-1
12	718.5	34.6	338	10	US-09-770-693-1
13	718.5	34.6	338	10	US-09-766-348-1
14	246.5	11.9	1079	9	US-09-820-843A-20
15	246.5	11.9	484	9	US-09-820-843A-19
16	220.5	10.6	651	10	US-09-861-597-1
17	211.5	10.2	606	10	US-09-861-597-6
18	198	9.5	344	9	US-10-034-158-7
19	197.5	9.5	344	9	US-10-034-158-7

20	197.5	9.5	344	10	US-09-086-118-27	Sequence 27, Appl
21	197.5	9.5	344	10	US-09-835-684-11	Sequence 11, Appl
22	197.5	9.5	344	10	US-09-880-371-11	Sequence 11, Appl
23	197.5	9.5	344	10	US-09-879-248-15	Sequence 15, Appl
24	197.5	9.5	344	10	US-09-770-693-7	Sequence 7, Appl
25	197.5	9.5	344	10	US-09-766-348-7	Sequence 7, Appl
26	197.5	9.5	357	10	US-09-864-761-35807	Sequence 35807, A
27	197.5	9.5	606	10	US-09-861-597-8	Sequence 8, Appl
28	192.5	9.3	606	10	US-09-861-597-4	Sequence 4, Appl
29	188	9.0	253	10	US-09-864-761-36720	Sequence 36720, A
30	187	9.0	256	9	US-09-820-843A-18	Sequence 18, Appl
31	175.5	8.4	611	9	US-10-086-510-4	Sequence 4, Appl
32	174.5	8.4	440	9	US-10-066-500-106	Sequence 106, Appl
33	174.5	8.4	440	9	US-10-063-547-52	Sequence 52, Appl
34	174.5	8.4	440	9	US-10-174-590-202	Sequence 202, Appl
35	174.5	8.4	440	9	US-10-176-758-202	Sequence 202, Appl
36	174.5	8.4	440	9	US-10-063-616-52	Sequence 52, Appl
37	174.5	8.4	440	9	US-10-175-737-202	Sequence 202, Appl
38	174.5	8.4	440	9	US-10-063-502-52	Sequence 52, Appl
39	174.5	8.4	440	9	US-10-173-706-202	Sequence 202, Appl
40	174.5	8.4	440	9	US-10-175-738-202	Sequence 202, Appl
41	174.5	8.4	440	9	US-10-175-752-202	Sequence 202, Appl
42	174.5	8.4	440	9	US-10-176-482-202	Sequence 202, Appl
43	174.5	8.4	440	9	US-10-176-757-202	Sequence 202, Appl
44	174.5	8.4	440	9	US-10-176-913-202	Sequence 202, Appl
45	174.5	8.4	440	9	US-10-180-552-202	Sequence 202, Appl

ALIGNMENTS

RESULT 1									
US-10-034-158-3									
Sequence 3, Appl									
Publication No. US20030028918A1									
GENERAL INFORMATION:									
APPLICANT: Wei, Zhong-Min									
TITLE OF INVENTION: METHOD OF IMPARTING DROUGHT RESISTANCE TO PLANTS									
FILE REFERENCE: 21839/230									
CURRENT APPLICATION NUMBER: US/10/034,158									
CURRENT FILING DATE: 2001-12-20									
PRIOR APPLICATION NUMBER: 09/597,840									
PRIOR FILING DATE: 2000-06-20									
PRIOR APPLICATION NUMBER: 09/013,587									
PRIOR FILING DATE: 1998-01-26									
PRIOR APPLICATION NUMBER: 60/036,048									
PRIOR FILING DATE: 1997-01-27									
NUMBER OF SEQ ID NOS: 10									
SOFTWARE: PatentIn Ver. 2.1									
SEQ ID NO 3									
LENGTH: 403									
TYPE: PRT									
ORGANISM: Erwinia amylovora									
US-10-034-158-3									
Query Match									
Best Local Similarity 100.0%; Pred. No. 9.5e-149; Length 403;									
Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MSLSITGSLASTMQLSISGA	NGGLGTSR	NAAGLGSALG	LGAGGNDV	YNOLAGLL	60		
DB	1	MSLSITGSLASTMQLSISGA	NGGLGTSR	NAAGLGSALG	LGAGGNDV	YNOLAGLL	60		
QY	61	TGMAAMAMGCGGGLGGLG	SGGLGSGGLG	BGLSNALNDML	GGSLNTLGS	RKGN	120		
DB	61	TGMAAMAMGCGGGLGGLG	SGGLGSGGLG	BGLSNALNDML	GGSLNTLGS	RKGN	120		
QY	121	TTSTTSPDLAGINSTGND	STGCTSTSS	SPMOOLAKMSE	ETIMQSLF	FGDCDGT	180		
DB	121	TTSTTSPDLAGINSTGND	STGCTSTSS	SPMOOLAKMSE	ETIMQSLF	FGDCDGT	180		
QY	181	OGSSSGKQPTGEONAKK	GVTDALSG	LMGNGISOLL	NGGIGGGG	GGNAGG	GLDGSSTL	240	
DB	181	OGSSSGKQPTGEONAKK	GVTDALSG	LMGNGISOLL	NGGIGGGG	GGNAGG	GLDGSSTL	240	

Db 181 QGSSSGGKOPTGEGBNAYKKGVTDALSGLMNGLSQLLGNGLGGGCGGAGNAGTGLDSSL 240
QY 241 GKGGLQNLGSPVDYQOLGNVGTGIGKAGIQALNDIGTHRSSTRFVNGDRAMAKEY 300
Db 241 GKGGLQNLGSPVDYQOLGNVGTGIGKAGIQALNDIGTHRSSTRFVNGDRAMAKEY 300
QY 301 GGFMDQYPEVFGKPOYKGPQGEVKTDDKSWAKALSRPDDGKMPASMEQFNKAKGMIKR 360
Db 301 GGFMDQYPEVFGKPOYKGPQGEVKTDDKSWAKALSRPDDGKMPASMEQFNKAKGMIKR 360
QY 361 PMAGDTGNGNLQARAGAGSSISGIDAMAGDAINNMALGKLGAA 403
Db 361 PMAGDTGNGNLQARAGAGSSISGIDAMAGDAINNMALGKLGAA 403

RESULT 2

US-09-086-118-23
Sequence 23, Application US/09086118
Patent No. US20010011380A1
GENERAL INFORMATION:
APPLICANT: Lady, Ronald J.
APPLICANT: Beer, Steven V.
APPLICANT: Wei, Zhong-Min
TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR
TITLE OF INVENTION: FRAGMENT ELICITING A HYPERSENSITIVE RESPONSE AND USES
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/086,118
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/048,109
FILING DATE: 30-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/1301
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 403 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-086-118-23

Query Match 100.0%; Score 2079; DB 10; Length 403;
Best Local Similarity 100.0%; Pred. No. 9,5e-149;
Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLNTSGIGASTMOISIGAGGNNGLGTSRONAGLGNSALGIGGNNQNDYVQNLGGL 60
Db 1 MSLNTSGIGASTMOISIGAGGNNGLGTSRONAGLGNSALGIGGNNQNDYVQNLGGL 60
QY 61 TGNMNMMSMNGGGGGLMGGLGGIGLGNGLGSSGGIGEGLSNALNDMLGSLNTLGSKGGNN 120
Db 61 TGNMNMMSMNGGGGGLMGGLGGIGLGNGLGSSGGIGEGLSNALNDMLGSLNTLGSKGGNN 120

QY 121 TTSTNSPLDQALGINSTQNDSDSTGTSDDSPMOQLLKRFSEIMQSLFGDGQGT 180
Db 121 TTSTNSPLDQALGINSTQNDSDSTGTSDDSPMOQLLKRFSEIMQSLFGDGQGT 180
QY 181 QGSSSGGKOPTGEGBNAYKKGVTDALSGLMNGLSQLLGNGLGGGCGGAGNAGTGLDSSL 240
Db 181 QGSSSGGKOPTGEGBNAYKKGVTDALSGLMNGLSQLLGNGLGGGCGGAGNAGTGLDSSL 240
QY 241 GKGGLQNLGSPVDYQOLGNVGTGIGKAGIQALNDIGTHRSSTRFVNGDRAMAKEY 300
Db 241 GKGGLQNLGSPVDYQOLGNVGTGIGKAGIQALNDIGTHRSSTRFVNGDRAMAKEY 300
QY 301 GGFMDQYPEVFGKPOYKGPQGEVKTDDKSWAKALSRPDDGKMPASMEQFNKAKGMIKR 360
Db 301 GGFMDQYPEVFGKPOYKGPQGEVKTDDKSWAKALSRPDDGKMPASMEQFNKAKGMIKR 360
QY 361 PMAGDTGNGNLQARAGAGSSISGIDAMAGDAINNMALGKLGAA 403
Db 361 PMAGDTGNGNLQARAGAGSSISGIDAMAGDAINNMALGKLGAA 403

RESULT 3

US-09-835-684-3
Sequence 3, Application US/09835684
Patent No. US20020019337A1
GENERAL INFORMATION:
APPLICANT: Qiu, Zhong-Min
APPLICANT: Remick, Dean
TITLE OF INVENTION: TREATMENT OF FRUITS OR VEGETABLES WITH HYPERSENSITIVE
TITLE OF INVENTION: RESPONSE ELICITOR TO CONTROL POSTHARVEST DISEASE OR
FILE REFERENCE: 21829/71
CURRENT APPLICATION NUMBER: US/09/835,684
CURRENT FILING DATE: 2001-04-16
PRIOR APPLICATION NUMBER: 60/198,359
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 403
TYPE: PRT
ORGANISM: Erwinia amylovora
US-09-835-684-3

Query Match 100.0%; Score 2079; DB 10; Length 403;
Best Local Similarity 100.0%; Pred. No. 9,5e-149;
Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLNTSGIGASTMOISIGAGGNNGLGTSRONAGLGNSALGIGGNNQNDYVQNLGGL 60
Db 1 MSLNTSGIGASTMOISIGAGGNNGLGTSRONAGLGNSALGIGGNNQNDYVQNLGGL 60
QY 61 TGNMNMMSMNGGGGGLMGGLGGIGLGNGLGSSGGIGEGLSNALNDMLGSLNTLGSKGGNN 120
Db 61 TGNMNMMSMNGGGGGLMGGLGGIGLGNGLGSSGGIGEGLSNALNDMLGSLNTLGSKGGNN 120
QY 121 TTSTNSPLDQALGINSTQNDSDSTGTSDDSPMOQLLKRFSEIMQSLFGDGQGT 180
Db 121 TTSTNSPLDQALGINSTQNDSDSTGTSDDSPMOQLLKRFSEIMQSLFGDGQGT 180
QY 181 QGSSSGGKOPTGEGBNAYKKGVTDALSGLMNGLSQLLGNGLGGGCGGAGNAGTGLDSSL 240
Db 181 QGSSSGGKOPTGEGBNAYKKGVTDALSGLMNGLSQLLGNGLGGGCGGAGNAGTGLDSSL 240
QY 241 GKGGLQNLGSPVDYQOLGNVGTGIGKAGIQALNDIGTHRSSTRFVNGDRAMAKEY 300
Db 241 GKGGLQNLGSPVDYQOLGNVGTGIGKAGIQALNDIGTHRSSTRFVNGDRAMAKEY 300
QY 301 GGFMDQYPEVFGKPOYKGPQGEVKTDDKSWAKALSRPDDGKMPASMEQFNKAKGMIKR 360
Db 301 GGFMDQYPEVFGKPOYKGPQGEVKTDDKSWAKALSRPDDGKMPASMEQFNKAKGMIKR 360

QY 361 PMAGDTGNGNLQARAGAGSSSLGIDAMMAGDAINNMALGKLGAA 403
 Db 361 PMAGDTGNGNLQARAGAGSSSLGIDAMMAGDAINNMALGKLGAA 403

RESULT 4

US-09-880-371-3
 ; Sequence 3, Application US/09880371
 ; Patent No. US20020059658A1
 ; GENERAL INFORMATION:

APPLICANT: Wei, Zhong-Min
 APPLICANT: Desocher, Jay
 TITLE OF INVENTION: METHODS OF IMPROVING THE EFFECTIVENESS OF TRANSGENIC
 FILE REFERENCE: 21829/91
 CURRENT APPLICATION NUMBER: US/09/880,371
 PRIOR FILING DATE: 2001-06-13
 PRIOR APPLICATION NUMBER: 60/211,585
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 3
 LENGTH: 403
 TYPE: PRT
 ORGANISM: Erythrina amylovora

Query Match

Best Local Similarity 100.0%; Score 2079; DB 10; Length 403;
 Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSINTSGIGASTMQLSISGAGAGNNGILGTSRQNGAGLSALGIGGQNDVTNQLAGLL 60
 Db 1 MSINTSGIGASTMQLSISGAGAGNNGILGTSRQNGAGLSALGIGGQNDVTNQLAGLL 60
 QY 61 TGMAMMAMGAGGGLMGGLGGLGNGIGSGGGLGEGLSNLMNMLGSLNTLGSKGGNN 120
 Db 61 TGMAMMAMGAGGGLMGGLGGLGNGIGSGGGLGEGLSNLMNMLGSLNTLGSKGGNN 120
 QY 121 TTSTNSPLDQALGINSTSQNDSTSGTSTSDSPMQQLLKPFSEIMOSLFGDGQDGT 180
 Db 121 TTSTNSPLDQALGINSTSQNDSTSGTSTSDSPMQQLLKPFSEIMOSLFGDGQDGT 180
 QY 181 QGSSSGKQPTGEBQNAKKGVTALSGIMNGLSQLLGNGLGGGCGNAGTGLDGSSL 240
 Db 181 QGSSSGKQPTGEBQNAKKGVTALSGIMNGLSQLLGNGLGGGCGNAGTGLDGSSL 240
 QY 241 GKGGLQNLGSPVDYQQLGNVGTGIGMKAGIOLNDIGTHHSSSTRSFVNGDRAMAKEI 300
 Db 241 GKGGLQNLGSPVDYQQLGNVGTGIGMKAGIOLNDIGTHHSSSTRSFVNGDRAMAKEI 300
 QY 301 GQFMDQYEVFGKQYQKGPQEVKTDKSNKALSKRPDDGMPASMEQFNKAKGMIKR 360
 Db 301 GQFMDQYEVFGKQYQKGPQEVKTDKSNKALSKRPDDGMPASMEQFNKAKGMIKR 360
 QY 361 PMAGDTGNGNLQARAGAGSSSLGIDAMMAGDAINNMALGKLGAA 403
 Db 361 PMAGDTGNGNLQARAGAGSSSLGIDAMMAGDAINNMALGKLGAA 403

RESULT 5

US-09-879-248-3
 ; Sequence 3, Application US/09879248
 ; Patent No. US20020062500A1
 ; GENERAL INFORMATION:

APPLICANT: Fan, Hao
 APPLICANT: Wei, Zhong-Min
 TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITING DOMAINS AND USE
 FILE REFERENCE: 21829/81
 CURRENT APPLICATION NUMBER: US/09/879,248
 CURRENT FILING DATE: 2001-06-12

PRIOR APPLICATION NUMBER: 60/212,211
 PRIOR FILING DATE: 2000-06-16
 NUMBER OF SEQ ID NOS: 18
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 3
 LENGTH: 403
 TYPE: PRT
 ORGANISM: Erythrina amylovora

Query Match

Best Local Similarity 100.0%; Score 2079; DB 10; Length 403;
 Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSINTSGIGASTMQLSISGAGAGNNGILGTSRQNGAGLSALGIGGQNDVTNQLAGLL 60
 Db 1 MSINTSGIGASTMQLSISGAGAGNNGILGTSRQNGAGLSALGIGGQNDVTNQLAGLL 60
 QY 61 TGMAMMAMGAGGGLMGGLGGLGNGIGSGGGLGEGLSNLMNMLGSLNTLGSKGGNN 120
 Db 61 TGMAMMAMGAGGGLMGGLGGLGNGIGSGGGLGEGLSNLMNMLGSLNTLGSKGGNN 120
 QY 121 TTSTNSPLDQALGINSTSQNDSTSGTSTSDSPMQQLLKPFSEIMOSLFGDGQDGT 180
 Db 121 TTSTNSPLDQALGINSTSQNDSTSGTSTSDSPMQQLLKPFSEIMOSLFGDGQDGT 180
 QY 181 QGSSSGKQPTGEBQNAKKGVTALSGIMNGLSQLLGNGLGGGCGNAGTGLDGSSL 240
 Db 181 QGSSSGKQPTGEBQNAKKGVTALSGIMNGLSQLLGNGLGGGCGNAGTGLDGSSL 240
 QY 241 GKGGLQNLGSPVDYQQLGNVGTGIGMKAGIOLNDIGTHHSSSTRSFVNGDRAMAKEI 300
 Db 241 GKGGLQNLGSPVDYQQLGNVGTGIGMKAGIOLNDIGTHHSSSTRSFVNGDRAMAKEI 300
 QY 301 GQFMDQYEVFGKQYQKGPQEVKTDKSNKALSKRPDDGMPASMEQFNKAKGMIKR 360
 Db 301 GQFMDQYEVFGKQYQKGPQEVKTDKSNKALSKRPDDGMPASMEQFNKAKGMIKR 360
 QY 361 PMAGDTGNGNLQARAGAGSSSLGIDAMMAGDAINNMALGKLGAA 403
 Db 361 PMAGDTGNGNLQARAGAGSSSLGIDAMMAGDAINNMALGKLGAA 403

RESULT 6

US-09-770-693-3
 ; Sequence 3, Application US/09770693
 ; Patent No. US20020069434A1
 ; GENERAL INFORMATION:

APPLICANT: Beer, Steven V.
 APPLICANT: Bauer, David W.
 TITLE OF INVENTION: COMBATE-RESISTANT TRANSGENIC PLANTS BY VIRTUE OF
 TITLE OF INVENTION: PATHOGEN-INDUCED EXPRESSION OF A HETEROLOGOUS
 FILE REFERENCE: 19603/2501
 CURRENT APPLICATION NUMBER: US/09/770,693
 PRIOR FILING DATE: 2001-01-26
 PRIOR APPLICATION NUMBER: 60/178,565
 NUMBER OF SEQ ID NOS: 26
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 3
 LENGTH: 403
 TYPE: PRT
 ORGANISM: Erythrina amylovora

Query Match

Best Local Similarity 100.0%; Score 2079; DB 10; Length 403;
 Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSINTSGIGASTMQLSISGAGAGNNGILGTSRQNGAGLSALGIGGQNDVTNQLAGLL 60
 Db 1 MSINTSGIGASTMQLSISGAGAGNNGILGTSRQNGAGLSALGIGGQNDVTNQLAGLL 60


```

RESULT 7
US-09-766-348-3
: Sequence 3, Application US/09766348
: Patent No. US2002011673A1
: GENERAL INFORMATION:
: APPLICANT: Qiu, Dewen
: APPLICANT: Wei, Zhong-Min
: APPLICANT: Beer, Steven V.
: TITLE OF INVENTION: HYPERSENSITIVE RESPONSE INDUCED RESISTANCE IN PLANTS BY
: TITLE OF INVENTION: SEED TREATMENT
: FILE REFERENCE: 19603/2986
: CURRENT APPLICATION NUMBER: US/09/766,348
: CURRENT FILING DATE: 2001-01-19
: PRIOR APPLICATION NUMBER: 08/984,207
: PRIOR FILING DATE: 1997-12-03
: PRIOR APPLICATION NUMBER: 60/033,230
: PRIOR FILING DATE: 1996-12-05
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 3
: LENGTH: 403
: TYPE: PRF
: ORGANISM: Erynia amylovora
: US-09-766-348-3

```

```

Db      241 GGGKGLONLSGVDPYQQLCNNAVGTGICMGKAG:IOALNDIGTTHRSSRSPRVNNGDRAMAEI 300
Qy      301 GGFMDQYPEVFGKROYOKPGQEVKTTDKSRNAKALSKRDDGGTTPASHEQNRKAKGMITR 360
Db      301 GGFMDQYPEVFGKROYOKPGQEVKTTDKSRNAKALSKRDDGGTTPASHEQNRKAKGMITR 360
Qy      361 PMAGDTGNGNLOARGAGSSSGIDPAMAGDAIINNALGKLGAA 403
Db      361 PMAGDTGNGNLOARGAGSSSGIDPAMAGDAIINNALGKLGAA 403

RESULT 8
US-10-034-158-1
: Sequence 1, Application US/10034158
: Publication No. US20030028918A1
: GENERAL INFORMATION:
: APPLICANT: Wei, Zhong-Min
: TITLE OF INVENTION: METHOD OF IMPARTING DROUGHT RESISTANCE TO PLANTS
: FILE REFERENCE: 21829/230
: CURRENT APPLICATION NUMBER: US/10/034, 158
: CURRENT FILING DATE: 2001-12-20
: PRIOR APPLICATION NUMBER: 09/597, 840
: PRIOR FILING DATE: 2000-06-20
: PRIOR APPLICATION NUMBER: 09/013, 587
: PRIOR FILING DATE: 1998-01-26
: PRIOR APPLICATION NUMBER: 60/036, 048
: PRIOR FILING DATE: 1997-01-27
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 338
: TYPE: PRT
: ORGANISM: Erynia chrysanthemi
US-10-034-158-1

Query Match      34.6%; Score 718.5; DB 9; Length 338;
Match Local Similarity 42.8%; Pred. No. 1.4e-46;
Matches 173; Conservatve 41; Mismatches 111; Indels 79; Gaps 11;

Qy      13 MOISI-GGAGGNNGLTGRONAGIGG-NSA---LGAGGNOGNDTVNOLAGITGMMAMM 67
Db      1 MGIITKAHIGGDDLVSGIAGQ--GIKGLNSAASSLGSSVVDLSTIIDKLTSLTSM--- 55
Qy      68 SMAGGGGIGGGGGLGAGLNGLGSSGGGLEGSLNALNDMLGSLNTLGSKGGNNTTSTNS 127
Db      56 -----FGALAAOGLGAS-SKELGMSNOIGOSFGN-----GAQGASNLLSYPK- 96
Qy      128 PLDQALGINSTQNDSDTSGTDTSDSDSPMOQLKMPSEIMOSLFG-----DGQDG 179
Db      97 -----SGGDALS-----KMFDKALDDLIGHDPTVKLTNOSNQ 128
Qy      180 TGGSSSGGKOPTGEGBONAKKGVTDALSGLGNGL:SOILGNGGLGGGCGGAGNAGTGLDSS 239
Db      129 LANMLNLAQYTOGNMNAFGSGVNNALSSII:GNGIGQSM-----SGFSQPS 174
Qy      240 LGGKGLNLSFPVYQOGLGNVVGICGKAKIOALNDIGTTHRSSRSPRVNNGDRAMAEI 299
Db      175 LGAGGGLGSLGAGFNOLGNAIGMGVGNAL:SLNSVSTHYHDGNNRFFVDEKRGMAKE 234
Qy      300 IGFMDQYPEVFGKROYOKPGQEVKTTDKSRNAKALSKRDDGGTTPASHEQNRKAKGMITR 359
Db      235 IGFMDQYPEVFGKROYOKPGQEVKTTDKSRNAKALSKRDDGGTTPASHEQNRKAKGMITR 294
Qy      360 RPAAGDTGNGNLOARGAGSSSGIDPAMAGDAIINNALGKLGAA 403
Db      295 SAVAGDTGNGNMLNRGAGGASIGIDPAAYVGDRIANMISIGKLGAA 338

RESULT 9
US-09-086-118-21
: Sequence 21, Application US/09086118
: Patent No. US20010011380A1
: GENERAL INFORMATION:

```

```

APPLICANT: Lady, Ronald J.
APPLICANT: Beer, Steven V.
TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR
TITLE OF INVENTION: FRAGMENT ELICITING A HYPERSENSITIVE RESPONSE AND USES
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hairgrave, Devans & Doyle LLP
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/086,118
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/048,109
FILING DATE: 30-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/1301
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 338 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-086-118-21

Query Match          34.6%; Score 718.5; DB 10; Length 338;
Best Local Similarity 42.8%; Pred. No. 1.4e-46;
Matches 173; Conservative 41; Mismatches 111; Indels 79; Gaps 11;

QY 13 MOISI-GGAGGNGNLLGTSRQNALGG-NSA---LGLGGGNDNTVQNLGLTGMMMM 67
  |||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|
  1 MQITKAHIGGDLGVSGLAGQ--GLKGLNSAASSIGSVKLSSTIDKLSALTSMM--- 55
QY 68 SMMGGGGLMGGGGLGGGGLGNGIGGEGISNLANLMDGLSLNTLSKSGGNNTTSTNS 127
  |||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|
  56 -----FGALAGLGAS-SKGLGMSNOLGSGFGN-----GAQASNLISVPK- 96
QY 128 PLDQALGINSTQNDSTSGTSDTSDDPMQQLKFESEIMOSLFG-----DGODG 179
  |||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|
  97 -----SGGDALS-----KMFDRALDDLGHDTYTKLTNSNQ 128
QY 180 TQSSSGGKOPTGEBQNAKRYGVTDALSGLMGNLSQLLNGGIGGCGGAGTGLDSS 239
  |||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|
  129 LANSMLNASQMTQGNNAFGSGVNNALSLTGLNGLSM-----SGFSQPS 174
QY 240 LGKGLQNLSPVYDQOLGNAVGTGIGMKAGIQALNDIGTHRSSTRESFVNKGDRAMAK 299
  |||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|
  175 LGAGGLOGLSGAGAFNOLGNALIGVGNALSLSVSTHVDGNHNFVDKEDRGAKE 234
QY 300 IGFPMQDYPEVFGKPOYOKGPGQEVKTDKSMARALSKPDDDCMTPASMEQFNKAKMIK 359
  |||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|
  235 IGFPMQDYPEVFGKPEYOKGQWSSPKTDDKSWAALSKPDDDCMTGASMDKFRAMKMIK 294
QY 360 RPAAGDTGNGNLQARGAGSSSLGIDAMAGDAINNMALGKIGAA 403
  |||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|
  295 SAVAGDTGNTNMLNRGAGASLIGDAVYVDKIANMSIGKIGAA 338
  |||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|
  DB

```

```

RESULT 10
US-09-835-684-1
Sequence 1, Application US/09835684
Patent No. US20020019337A1
GENERAL INFORMATION:
APPLICANT: Wei, Zhong-Min
APPLICANT: Qiu, Dewen
APPLICANT: Remick, Dean
TITLE OF INVENTION: TREATMENT OF FRUITS OR VEGETABLES WITH HYPERSENSITIVE
TITLE OF INVENTION: RESPONSE ELICITOR TO CONTROL POSTHARVEST DISEASE OR
TITLE OF INVENTION: DESICCATION
FILE REFERENCE: 21829/71
CURRENT APPLICATION NUMBER: US/09/835,684
CURRENT FILING DATE: 2001-04-16
PRIOR APPLICATION NUMBER: 60/198,359
PRIOR FILING DATE: 2000-04-19
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 338
TYPE: PRT
ORGANISM: Erythrina chrysanthemi
US-09-835-684-1

Query Match          34.6%; Score 718.5; DB 10; Length 338;
Best Local Similarity 42.8%; Pred. No. 1.4e-46;
Matches 173; Conservative 41; Mismatches 111; Indels 79; Gaps 11;

QY 13 MOISI-GGAGGNGNLLGTSRQNALGG-NSA---LGLGGGNDNTVQNLGLTGMMMM 67
  |||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|
  1 MQITKAHIGGDLGVSGLAGQ--GLKGLNSAASSIGSVKLSSTIDKLSALTSMM--- 55
QY 68 SMMGGGGLMGGGGLGGGGLGNGIGGEGISNLANLMDGLSLNTLSKSGGNNTTSTNS 127
  |||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|
  56 -----FGALAGLGAS-SKGLGMSNOLGSGFGN-----GAQASNLISVPK- 96
QY 128 PLDQALGINSTQNDSTSGTSDTSDDPMQQLKFESEIMOSLFG-----DGODG 179
  |||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|
  97 -----SGGDALS-----KMFDRALDDLGHDTYTKLTNSNQ 128
QY 180 TQSSSGGKOPTGEBQNAKRYGVTDALSGLMGNLSQLLNGGIGGCGGAGTGLDSS 239
  |||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|
  129 LANSMLNASQMTQGNNAFGSGVNNALSLTGLNGLSM-----SGFSQPS 174
QY 240 LGKGLQNLSPVYDQOLGNAVGTGIGMKAGIQALNDIGTHRSSTRESFVNKGDRAMAK 299
  |||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|
  175 LGAGGLOGLSGAGAFNOLGNALIGVGNALSLSVSTHVDGNHNFVDKEDRGAKE 234
QY 300 IGFPMQDYPEVFGKPOYOKGPGQEVKTDKSMARALSKPDDDCMTPASMEQFNKAKMIK 359
  |||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|
  235 IGFPMQDYPEVFGKPEYOKGQWSSPKTDDKSWAALSKPDDDCMTGASMDKFRAMKMIK 294
QY 360 RPAAGDTGNGNLQARGAGSSSLGIDAMAGDAINNMALGKIGAA 403
  |||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|
  295 SAVAGDTGNTNMLNRGAGASLIGDAVYVDKIANMSIGKIGAA 338
  |||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|  ||:|
  DB

RESULT 11
US-09-880-371-1
Sequence 1, Application US/09880371
Patent No. US20020059658A1
GENERAL INFORMATION:
APPLICANT: Wei, Zhong-Min
APPLICANT: DeRoche, Jay
TITLE OF INVENTION: METHODS OF IMPROVING THE EFFECTIVENESS OF TRANSGENIC
TITLE OF INVENTION: PLANTS
FILE REFERENCE: 21829/91
CURRENT APPLICATION NUMBER: US/09/880,371
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: 60/211,585
PRIOR FILING DATE: 2000-06-15

```

NUMBER OF SEQ ID NOS: 16
 SOFTWARE: Patentin Ver. 2.1
 SEQ ID NO 1
 LENGTH: 338
 TYPE: PRT
 ORGANISM: Erwinia chrysanthemi
 US-09-880-371-1

Query Match 34.6%; Score 718.5; DB 10; Length 338;
 Best Local Similarity 42.8%; Pred. No. 1.4e-46;
 Matches 173; Conservative 41; Mismatches 111; Indels 79; Gaps 11;

QY 13 MOISI-GGAGGNNGLTSTRONAGLIG-NSA---LGLGGGNQNDTVNQLAGLTGMMMM 67
 DB 1 MOTIRAHIGDGLVSSGLGAG--GLKGLNSAASSLGSVDKLSSTIDKLTSLTSM-- 55
 QY 68 SMGGGGLMGGGLGGLGNGLIGSGGGLGEGLSNLMNLMGLSLTLTSGSKGNNTTSTNS 127
 DB 56 -----FGALAGLIGAS-SKGLGMSNOLGSGFN-----GAGASNLISVPK- 96
 QY 128 PLDQALGINTSQNDSTSGTSDTSDSDPMQOLLKMFSEIMOSLFG-----DQDQ 179
 DB 97 -----SGGDALS-----KMFDRKALDLDLGHDTYTKLTNSQ 128
 QY 180 TCGSSSGKOPTGEQNAAYKKGYTDALSGLMGNGLSOLLGNGGGLGGGAGAGTGLDSS 239
 DB 129 LANSMLNASQMTQGNMNAFGSGVNNALSLILGNGLSGM-----SGFSQPS 174
 QY 240 LGGKGLONLSPVDYQOLGNAVGTGIGMKAGIQALNDIGTHRSSTRSFYNGKGRAMAKE 299
 DB 175 LGAGGLGGLSGAGAFNQLGNLIGMGVGNALSLNSVSTHYDGNHNFYDKEDRGMAKE 234
 QY 300 IGFMDQYPEYFGAPYQKPGQEVKTDDKSWAKALSKPDDGKTPASMDQFNKAKMIK 359
 DB 235 IGFMDQYPEYFGAPYQKPGQEVKTDDKSWAKALSKPDDGKTPASMDQFNKAKMIK 294
 QY 360 RPMAGDTGNGMLQARGAGSSSLGIDAMAGDAINNMLGKIGAA 403
 DB 295 SAVAGDTGNTNMLNRGAGGASLIGIDAAYVGDKIANMISGLKIANA 338

RESULT 12
 US-09-879-248-1
 Sequence 1, Application US/09879248
 Patent No. US20020062500A1
 GENERAL INFORMATION:
 APPLICANT: Fan, Hao
 TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITING DOMAINS AND USE
 FILE REFERENCE: 21829/81
 CURRENT APPLICATION NUMBER: US/09/879, 248
 PRIOR FILING DATE: 2001-06-12
 PRIOR APPLICATION NUMBER: 60/212,211
 NUMBER OF SEQ ID NOS: 18
 SOFTWARE: Patentin Ver. 2.1
 SEQ ID NO 1
 LENGTH: 338
 TYPE: PRT
 ORGANISM: Erwinia chrysanthemi
 US-09-879-248-1

Query Match 34.6%; Score 718.5; DB 10; Length 338;
 Best Local Similarity 42.8%; Pred. No. 1.4e-46;
 Matches 173; Conservative 41; Mismatches 111; Indels 79; Gaps 11;
 QY 13 MOISI-GGAGGNNGLTSTRONAGLIG-NSA---LGLGGGNQNDTVNQLAGLTGMMMM 67
 DB 1 MOTIRAHIGDGLVSSGLGAG--GLKGLNSAASSLGSVDKLSSTIDKLTSLTSM-- 55
 QY 68 SMGGGGLMGGGLGGLGNGLIGSGGGLGEGLSNLMNLMGLSLTLTSGSKGNNTTSTNS 127

DB 56 -----FGALAGLIGAS-SKGLGMSNOLGSGFN-----GAGASNLISVPK- 96
 QY 128 PLDQALGINTSQNDSTSGTSDTSDSDPMQOLLKMFSEIMOSLFG-----DQDQ 179
 DB 97 -----SGGDALS-----KMFDRKALDLDLGHDTYTKLTNSQ 128
 QY 180 TCGSSSGKOPTGEQNAAYKKGYTDALSGLMGNGLSOLLGNGGGLGGGAGAGTGLDSS 239
 DB 129 LANSMLNASQMTQGNMNAFGSGVNNALSLILGNGLSGM-----SGFSQPS 174
 QY 240 LGGKGLONLSPVDYQOLGNAVGTGIGMKAGIQALNDIGTHRSSTRSFYNGKGRAMAKE 299
 DB 175 LGAGGLGGLSGAGAFNQLGNLIGMGVGNALSLNSVSTHYDGNHNFYDKEDRGMAKE 234
 QY 300 IGFMDQYPEYFGAPYQKPGQEVKTDDKSWAKALSKPDDGKTPASMDQFNKAKMIK 359
 DB 235 IGFMDQYPEYFGAPYQKPGQEVKTDDKSWAKALSKPDDGKTPASMDQFNKAKMIK 294
 QY 360 RPMAGDTGNGMLQARGAGSSSLGIDAMAGDAINNMLGKIGAA 403
 DB 295 SAVAGDTGNTNMLNRGAGGASLIGIDAAYVGDKIANMISGLKIANA 338

RESULT 13
 US-09-770-693-1
 Sequence 1, Application US/09770693
 Patent No. US20020069434A1
 GENERAL INFORMATION:
 APPLICANT: Baer, Steven V.
 TITLE OF INVENTION: COMETEGE-RESISTANT TRANSGENIC PLANTS BY VIRTUE OF
 TITLE OF INVENTION: PATOEGE-INDUCED EXPRESSION OF A HETEROLOGOUS
 FILE REFERENCE: 19603/2501
 CURRENT APPLICATION NUMBER: US/09/770, 693
 PRIOR FILING DATE: 2001-01-26
 PRIOR APPLICATION NUMBER: 60/178,565
 NUMBER OF SEQ ID NOS: 26
 SOFTWARE: Patentin Ver. 2.1
 SEQ ID NO 1
 LENGTH: 338
 TYPE: PRT
 ORGANISM: Erwinia chrysanthemi
 US-09-770-693-1

Query Match 34.6%; Score 718.5; DB 10; Length 338;
 Best Local Similarity 42.8%; Pred. No. 1.4e-46;
 Matches 173; Conservative 41; Mismatches 111; Indels 79; Gaps 11;
 QY 13 MOISI-GGAGGNNGLTSTRONAGLIG-NSA---LGLGGGNQNDTVNQLAGLTGMMMM 67
 DB 1 MOTIRAHIGDGLVSSGLGAG--GLKGLNSAASSLGSVDKLSSTIDKLTSLTSM-- 55
 QY 68 SMGGGGLMGGGLGGLGNGLIGSGGGLGEGLSNLMNLMGLSLTLTSGSKGNNTTSTNS 127
 DB 56 -----FGALAGLIGAS-SKGLGMSNOLGSGFN-----GAGASNLISVPK- 96
 QY 128 PLDQALGINTSQNDSTSGTSDTSDSDPMQOLLKMFSEIMOSLFG-----DQDQ 179
 DB 97 -----SGGDALS-----KMFDRKALDLDLGHDTYTKLTNSQ 128
 QY 180 TCGSSSGKOPTGEQNAAYKKGYTDALSGLMGNGLSOLLGNGGGLGGGAGAGTGLDSS 239
 DB 129 LANSMLNASQMTQGNMNAFGSGVNNALSLILGNGLSGM-----SGFSQPS 174
 QY 240 LGGKGLONLSPVDYQOLGNAVGTGIGMKAGIQALNDIGTHRSSTRSFYNGKGRAMAKE 299
 DB 175 LGAGGLGGLSGAGAFNQLGNLIGMGVGNALSLNSVSTHYDGNHNFYDKEDRGMAKE 234
 QY 300 IGFMDQYPEYFGAPYQKPGQEVKTDDKSWAKALSKPDDGKTPASMDQFNKAKMIK 359
 DB 235 IGFMDQYPEYFGAPYQKPGQEVKTDDKSWAKALSKPDDGKTPASMDQFNKAKMIK 294

Thu Jun 12 14:55:50 2003

us-09-412-100-23.rapb

Page 7

```

QY      360 RPAAGDTGNGNLQARGAGSSSLGIDAMMAGDAINNMLGKLGAA 403
      : ||||| | | ||||:||||| : || | ||:||||
Db      295 SAVAGDTGNTNLNRGAGGASLGIDAAYVGRIANMSLGKLANA 338

```

```

RESULT 14
US-09-766-348-1
: Sequence 1, Application US/097666348
: Patent No. US20020116733A1
: GENERAL INFORMATION:
: APPLICANT: Qiu, Dewen
: APPLICANT: Wei, Zhong-Min
: APPLICANT: Beer, Steven V.
: TITLE OF INVENTION: HYPERSENSITIVE RESPONSE INDUCED RESISTANCE IN PLANTS BY
: TITLE OF INVENTION: SEED TREATMENT
: FILE REFERENCE: 19603/2986
: CURRENT APPLICATION NUMBER: US/09/766,348
: CURRENT FILING DATE: 2001-01-19
: PRIOR APPLICATION NUMBER: 08/984,207
: PRIOR FILING DATE: 1997-12-03
: PRIOR APPLICATION NUMBER: 60/033,230
: PRIOR FILING DATE: 1996-12-05
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 338
: TYPE: PRF
: ORGANISM: Erythrina chrysanthemi
: US-09-766-348-1

```

Query Match	34.6%;	Score 718.5;	DB 10;	Length 338;
Best Local Similarity	42.8%;	Pred. No. 1.4e-46;		
Matches 173;	Conservative 41;	Mismatches 111;	Indels 79;	Gaps 11;

[illegible]

```

RESULT 15
US-09-820-843A-20
: Sequence 20, Application US/09820843A
: Publication No. US20030039963A1
: GENERAL INFORMATION:
: APPLICANT: Council of Scientific and Industrial Research
: TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEIN
: TITLE OF INVENTION: USEFUL AS ANTI-INFECTIVES
: FILE REFERENCE: 063915

```

```

: CURRENT APPLICATION NUMBER: US/09/8320,843A
: CURRENT FILING DATE: 2001-03-30
: NUMBER OF SEQ ID NOS: 118
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 20
: LENGTH: 1079
: TYPE: PRT
: ORGANISM: M. tuberculosis
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: PE_PGSR
: NAME/KEY: misc_feature
: OTHER INFORMATION: g112924449
US-09-820-843A-20

```

Query Match	11.9%;	Score 246.5;	DB 9;	Length 1079;
Best Local Similarity	27.0%;	Pred. No. 1.7e-10;		
Matches 119;	Conservative 31;	Mismatches 168;	Indels 123;	Gaps 18;

Oy 5 TSGISLASTMISTIGGAGNNKLLGSTRONAGLGGSSALGIGGONOPDVNLAGLIGMM 64
 Db 636 TGGAGNGGGGAGNGAGGAGGAGGSGGGTGGNGGAGGDAAGAGNGMNGTGMGNG- 669
 Oy 65 MMSMMMGGL--WGGGLGGJLNGLGGSGGLGEGJ.SNALNDMLGSLNT---LGSKGN 119
 Db 690 -----GNGIAGMGNGGAGTGGGNGNGSGG--GNGAGNMGNGSGTSGGGAGGN 74.0
 Oy 120 NNTSTNPLDQALGINST-----SQMDSTSGTSDTSBSSDPMOQLMFSEIM 169
 Db 741 GGAAGGTGGTGDDG-GLTGTGTGGSGGTGGDGGNGGNADMTAN-----MT 785
 Oy 170 QSLFEGODDGTQSSSGGKQPTGEBQNAKKKVTDLGLGNGLSOLLGNGLG- 225
 Db 786 AQAAGGCGGCGGCGFGGAGAGGGGLTGANG--TGGGAGAGDG---GNGALGGGPL 839
 Oy 226 --GCGNAGTGLDGS--LGGKGLNLSGPV-----DYQJLNAVGTGIGM 267
 Db 840 TDDPGGCGGNGGCGGTGGAGIGSLGGTGGDGGNGNGGTGGEGGEGVAGAGT--GG 897
 Oy 268 KAGJQALNDIGTHRHSSSTRSFVKKGBRAAKKEIGFQMDQYEVFEGKPYQKGPQEVKTD 327
 Db 898 AAG-----NGDGGTGGTGG-----GDDGAGGTG 921
 Oy 328 DKSMKALSKP-----DDGKMTPASMEQPNFAKGMKIRPAGDGTGNGNTQARAGGSSL 381
 Db 922 GTGGTGLGDPFRVGGSGGCGGTGGSGGANGNG--GNAGAGCGNGNGTGTGAGG--I 975
 Oy 382 GIDAMAGDAILNNALGKIGA 402
 Db 976 GGTGGNGCGDAEPGVPAGGA 996

Search completed: June 12, 2003, 08:51:27
Job time : 24 secs

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-591 <COL>
A:Cross-references: GB:296800; GB:AL123456; NID:g3261800; PIDN:CAM09596.1; PID:e3216
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV0297
C:Superfamily: unassigned collagens

Query Match 12.3%; Score 255; DB 2; Length 591;
Best Local Similarity 26.8%; Pred. No. 7,6e-09;
Matches 123; Conservative 24; Mismatches 168; Indels 144; Gaps 20;

OY 7 GLGASTMOISTIGGAGCGNNGLGTSRONAGLGSNSALGLGCGNNDPTVQJLITGMMX 66
DB 141 GNGSGGAPGAGGAGGAAGFTG---NGNGGDCGAGANGG-----ACGTAGMFG 187
OY 67 MSMMGGGGLMG-----GGIGGGIGN-----GIGSGGIGBLSMALNDML----- 106
DB 188 FCGNGGAGGIGVAGINGLIGLGGAGGCGNAGPFGNGNGGAGAGAGAAVNAVNGLATPYT 247
OY 107 -----GSSLNTLGSKRG--GNNTTSTNSELDPALGINSTQNDSTSG-----TGSTSDS 154
DB 248 PAANGNGLINTVGVPTGAGGADGANGSAIQAGAGGADGNASTSGGIGIQTGAGGA 307
OY 155 SDPMQQLKMFSEIMQSLFPGQGDPTGSSSGCKPTEBEONAYKKGYDALSLGNGNL 214
DB 308 G-----GAGGDGAPGNGNGSGVEIT-----GAAGS--SASGNGA 342
OY 215 SOLLGNGLIG-----GCGGNGAGTGLDSSSLGKGLONSGPV 252
DB 343 TG--GNGGVGAPGAGGNGGHHVSGGSVMTAGAGGKRGNGCTGGAGP--GHRGSLSGPV 399
OY 253 DYQDLGNVAVGIGKAGIQLNDYIGTHRHSTRS----FVKGDRMAKETIGQFMDQYR 308
DB 400 GDGSGNGAGGDS--GAGSATDIAGTGGRGNGGNGHGLIMINGGDDGCGAGVG-- 449
OY 309 EVFGKPYOKGCGQGVETDDKSMAKALSKPDDDGMTPASMGQFNKAKGMIKRP--AG 364
DB 450 -----GVG-----GAGACALIGHGDDG-----GSVNTPIGSGEAG 480
OY 365 DTGNGNTLARGAGGSSSLGIDAMAGDAINNMALGKLGAA 403
DB 481 DGGKGLGDDGGGRGIFG--DFGAGGAGGAGGAVGAGGA 517

RESULT 4
C70974
hypothetical glycine-rich protein RV3388 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #extl_change 20-Jun-2000
C:Accession: C70974
R:Colo, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, R.A.; Rogers, J.; Devlin, K.; Felkell, T.; Gentles, S.; Hamlin, N.; Holtroyd, R.; Andrews, M.A.; Davies, R.; Kutter, S.; Seeger, K.; Skellton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MIDB:98295987; PMID:9634230
A:Accession: C70974
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-731 <COL>
A:Cross-references: GB:AL009198; GB:AL123456; NID:g3242262; PIDN:CA15773.1; PID:g2626
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV3388
C:Superfamily: elastin

Query Match 12.0%; Score 249; DB 2; Length 731;
Best Local Similarity 28.7%; Pred. No. 2.3e-08;
Matches 96; Conservative 23; Mismatches 101; Indels 114; Gaps 15;

A:Accession: F70971
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-588 <Col>
A:Cross-references: GB:AL009199; GB:AL123456; NID:93242262; PIDN:CAA15752.1; PID:e120222
A:Experimental source: strain H37Kv
A:Genetics:
A:Gene: RV3367
A:Superfamily: unassigned collagens

```

Query Match 11.9% Score 247; DB 2: Length 588;
Best Local Similarity 30.4% Pred. No. 2.4e-08;
Matches 117; Conservative 16; Mismatches 154; Indels 98; Gaps 21;

QY 18 GGAGC--NCILLGSRONAGICGNSALGLGCGNONTVYNQLAGLLTGMMMSWGCGGL 75
   |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 283 GGAGCGDANGNPANPNTSIANGAGGNGAAG-GGGANGCAGCAGCGAASNGSSVGGDCGNG 341
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 76 MGG---GLGGGGLGNGJAGSGGGLGEGLSNMLNDMLGSLNLTGSKGGNNTTSTTNSPLDQ 131
   ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 342 AGCGGTGNGHMGAG-GAGGAGGRCGMLVG--NGNGGN---GAAGCGAIGGTG--- 390
   ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||

QY 132 ALGINTSRONDNDTSGTSTSDSDPMQQLKMFSEIMQSLFPGCGDCTGSSSGGQPT 191
   ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 391 AGGV-PANCGNSALGT-----QPVGGGSGGNGNKTGG--T 424
   ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||

QY 192 EGEOINAYKKVYDALSG-IMGNGLSOLLGNGGLG--GGCGNACTGLDSSLAGCKIQL 248
   ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 425 GGRGCGDGGGAGAGGAGCWMIMGNGNG--GNCGTGCGSGVGNGGIGGAG-GGNATSTS 481
   ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||

QY 249 SGPDVYQDLGAAVVTGIGMAGIQLANDIGTHRISSTRSFYNNKEDRAMAKTIGCFMDIYP 308
   ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 482 SIIPDAHG-GAG--GAGGDAG-----HGVTGGDGDGHHGTGRRGLT---- 522
   ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||

QY 309 EYFEGPOYOKPGQGEVETDTRSMARALKRPDDGCTPASMEOFKAKGMIRKPAQTGN 368
   ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 523 -----AGGHANSNGCGGGGTGGAGGTG-TRGSGN-----AGGTGT 557
   ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||

QY 369 GNLQARGAGGSSLGIDAMAGDAIN 393
   ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 558 GNAUSTNGRGSDG---LGGDAFN 578
   ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||

```

RESULT 8

B70807
hypothetical glycine-rich protein RV3512 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_Change 20-Jun-2000
C:Accession: B70807

R:Cole, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference numbers: A70500; MUID:96295987; PMID:9634230
A:Accession: B70807

A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1079 <COL>

A:Cross-references: GB:AL022022; GB:AL123456; MID:g3261554; PIDN:CAAI7749.1; PTD:g2924444
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV3512

C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology

Query Match 11.9%; Score 246.5; DB 2; Length 1079;
Best Local Similarity 27.0%; Pred. No. 5.3e+08;
Matches 119; Conservative 31; Mismatches 168; Indels 123; Gaps 18;

OY 5 TSGAGSMQMSIGGAGGNGLGTSRONAGLGNSALGLGGGNONTVNOIAGLTIGMM 64
| | | | | | | | | | | | | | | | | | | | | | | | | |
DB 636 TGAGAGNGCGANGAGGAGGGGTGTGNGGACDAGNGAGNGTGNGGNG----- 689

OY	65	MAMSMWGGGL--MGGGJGGGJGNLJGSGCGJGBELSNALMDMLGGSINT---	LGSXGCN	119
Dd	690	-----GNGJLAOMGNGAGTGTSGGNGGSSG---GGGAAGAAGNCSTGSSGDGAGGN	740	
OY	120	NNTSTNSPLDQALGINST-----SQUDDSTSGTSDTSPOSDPMQQLMFGEIM	169	
Dd	741	GGAAGTGCTGSDG-GLTGTGCTGTGSSGCTGGDGNGKRNADWTAN-----MT	785	
OY	170	QSLEFGDQDSTOGSSSGKOPTGEQONAYKKGVTDALSGLMGNCLSOLLNGSLG-	225	
Dd	786	AQAAGCDGNGNDGDFGFGAGAGGGGLTAGANG-TGCGGAGAGDG-----GNGALTGHGPL	839	
OY	226	--GQGNAAGCJLDSS--LGKGLNLNSPV-----DTQJLNAVGTETGM	267	
Dd	840	TDDPGGNGGTGGNGGTGTGAGATSLDGCTGTGDCGNGNGCTGCGEGCEVYGACGT--GG	897	
OY	268	KAGIOLALDIGTRHSSTRSPFNKGSRPAWAKEIGFMDQYPEVFGRKPQYOGPQCJEKTD	327	
Dd	898	AAG-----MGSDGTGGTGG-----GDBGAGCTG	921	
OY	328	DKSWAKALSKP-----DDGHTPASMEQFNFAKMIRPNAGDTGNGNTLARAGAGSSSL	381	
Dd	922	GTGTGTGLCPFRVGSGSGDGCTGSSGGAAGNGNG-----GNAGAGGNGNGTGTGAGG-I	975	
OY	382	GIDAWMAADAINMALRTIGA	402	
Dd	976	GGTGGNGGDAPEGPCPAGCA	996	

RESULT 9
 H70846
 hypothetical glycine-rich protein RV3345c - Mycobacterium tuberculosis (strain H37RV
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C:Accession: H70846
 R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordo
 J.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holtroyd
 R.; Jandrem, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete gen
 A:Reference number: AY0500; MUID:98295987; PMID:9634230
 A:Accession: H70846
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1538 <COL>
 A:Cross-references: GB:AL021841; GB:AL123456; NID:g3261517; PIDN:CAAL7117.1; PID:g32
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: RV3345c
 C:Superfamily: collagen alpha 1(IV) chain

Query Match	11.8%	Score 246;	DB 2;	Length 1538;
Best Local Similarity	27.2%	Pred. No. 8.5e-08;		
Matches 124;	Conservative 32;	Mismatches 168;	Indels 132;	Gaps 21;

OY	7	GLGASTM	DISIGAGGNNLLGTSRONAGLG	-----NSALGLGCGNQNDTVN	-----	54
		:				
Db	249	GIG-----	GIGMGDGLGF	----	NCGAGGAGAAGLPGAAGLNGCGDSDGNGCTG	298
OY	55	-QLAGLGN	MMMMNMGGGILMG	-----	GLGGGLGN-GLGSSGG	93
Db	299	GGRGLLVG----	NCGAGGAGCGVGGDGCKGAGDPSFVYNNAGCGNGSHGNPVGAGG	354		
OY	94	LGEGISNALNMLG	SGSLNTLIGSKGNNNTSTTNSLPDLAGINSTQDNDDSTSGTYS	153		
		:	:	:	:	
Db	355	AG-GILTAGHAAAGAT	PTPSGNGCGKGIGAFANSP	411		
		:	:	:	:	
OY	154	SSDPWQLLKKFSEIM	QSLFGDQDQCGNCGSSCGQPEBGRONAKKRYVTALSGIMNG	213		
		:	:	:	:	
Db	412	AG-----	-GAGHAGSGGATGTLPTGGNGT	450		
		:	:	:	:	

A:Cross-references: GB:280225; GB:AL123456; NID:g3261265; PIDN:CAB02341.1; PID:e266390;
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV2634c
C:Superfamily: unassigned collagens

Query Match 11.6%; Score 242; DB 2; Length 778;
Best Local Similarity 30.18; Pred. No. 6.9e-08;
Matches 91; Conservative 19; Mismatches 110; Indels 82; Gaps 14;

QY 7 GLGASTWQISIGAG--GNNGLGTSRONAGLGSNSALGIG--GNONDYVWQLGIL 60
DB 353 GNGGNMNMWSSGAGGCGCTGAGTNGVNGSTANPTGTANGNDNSGNGOT----- 404
QY 61 TGMAMMMWMMGGGLMGGLGGLG--NGIGSGGLGEGLSNMLNDMLGSLNTLTKRG 118
DB 405 -----GNGG--GPGPAGGVGAGVGCGGGLGESLDG--NDGTG-----KGG 443
QY 119 NNTSTTNSPLDQALGINTSNDSTSGTSTSDSSDPMOQLKMFSEIMOSLFGDGD 178
DB 444 AGGTAGTGDGAGAGAGGIGETDGSAGV-----ATGGEAGD 481
QY 179 GTGSSSGCKPTEGE-QNAYKKGYTDAL---SLMGNG---LSQLLNGGLG----- 225
DB 482 GATGCVDGVGAGGKGGCGHNTGVGDARFGDGLGEGALGAMGNGGTGAGGNGG 541
QY 226 -----GCGNAGTGLDSSLGKGLQNLSPV---DYQQLGNVGTGIGMKAGIQLND 276
DB 542 RGMGLGNGGAGGAGGTGT--GGGGAAGFAGVGAGGGBELTGTGAGTAEAGTGLGLG 600
QY 277 IG 278
DB 601 VG 602

RESULT 13

hypothetical glycine-rich protein Rv0278c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence.revision 17-Jul-1998 #text.change 20-Jun-2000
C:Accession: D70835
R:Coile, S.T.; Brosch, R.; Parish, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; M0ID:98295987; PMID:9634230
A:Accession: D70835
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-957 <COL>
A:Cross-references: GB:AL021930; GB:AL123456; NID:g3261524; PIDN:CA17353.1; PID:g290946
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv0278c
C:Superfamily: elastin

Query Match 11.6%; Score 241.5; DB 2; Length 957;
Best Local Similarity 25.8%; Pred. No. 9.4e-08;
Matches 117; Conservative 42; Mismatches 149; Indels 145; Gaps 21;

QY 2 SLTSTGLASTMOWISIGAGNGNLLGTSRONALGCG--NSALGIGGNGNDYVWQLA 57
DB 362 TLVAGTIGCA-----CGAGGNAGLLFGSGSGAGGAFADGCGGPGCN-----A 407
QY 58 GLLTGMAMMMWMMGGGLMG--GGLGGLGNGLGSSGL-----GGLSNAL---N 103
DB 408 G-----TVFGSGAGNGGVGGGAGGAGTGPGLGNGGNGGASAVYGCN 458
QY 104 DMLGSLNTLTKRGKGNNTTSTNSPLDQALGINTSNDSTSGTSTSDSSDPMOQLK 163
DB 459 GGIGGTGVLGN--GNGGSGGIGAKRAGVGSGLLGLDGFNAPASTSLHTLQDNVLA 517

QY 164 MFSEIMOS-----LFGDGODGTGGS-----SSGKQPIEGEBONAYKKGYTDALSG-LMGNG 213
DB 518 VVNEPPLQTLGRPLRIGNANTPTGTADGAGGGLFNGANGTGTCTGAGAGGATLNGC 577
QY 214 LSQLLNGGLGCG-----GCGNAGTGLDSSLGKGLQNLSPV 254
DB 578 -----GNGGCHATNTAATATGAGAGGILFEGTNGGTG--GIATGAGGIGAGGAGV 630
QY 255 QQLGNAYGTG-----IGMKAGIQLNDIGTHRSSTRSVYKCDRMAKEIGOFMDQYP 308
DB 631 SLTIGSGGTGNGNSTGV--AGIGAGGAGG-----GDAGL----- 664
QY 309 EVEGKPOYOKPGCEVETDCKSMNAKSKPDDDCMTPASMEQFKAQMKIRPAGPTGN 368
DB 665 LFG-----AGTGC-----HGAAGVPAGGAGGN 690
QY 369 GNLDARAGAGSSSLGIDMAMGDAINNALKRLG 401
DB 691 GGLFANGAGAGGAGFNA--AGNGNGGLFETGCG 722

RESULT 14

hypothetical glycine-rich protein Rv1840c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence.revision 17-Jul-1998 #text.change 20-Jun-2000
C:Accession: H70663
R:Coile, S.T.; Brosch, R.; Parish, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; M0ID:98295987; PMID:9634230
A:Accession: H70663
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-515 <COL>
A:Cross-references: GB:283859; GB:AL123456; NID:g3261678; PIDN:CAB06114.1; PID:g1781
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv1840c
C:Superfamily: Phaseolus glycine-rich cell wall protein 1.8

Query Match 11.5%; Score 238.5; DB 2; Length 515;
Best Local Similarity 27.38; Pred. No. 7.1e-08;
Matches 109; Conservative 23; Mismatches 128; Indels 139; Gaps 15;

QY 18 GGAGGNNGLGTSRONAGLGSNSALGIGGNGNDYVWQLGILLTGMAMMMWMMGGGLMG 77
DB 148 GGAGGAGGLGNG--GNGCIGGPAATGLAG-----AGVGGLLFGDGGNGAGGLG 197
QY 78 GGLGGLGNGLGSSGGLGEGLSNMLNDMLGSLNTLTKSGKGNNTTSTNSPL--DNLGI 135
DB 198 TGPVGATG--GIGGPGCAAVGLFCHGAGAGGLAKAGFAGAGGTGTGLLNGNGNG 256
QY 136 NSTSONDSTSGTSTSDSSDPMOQLKMFSEIMOSLFGGCGGTGCGSSSGGKQPTGEGQ 195
DB 257 NVBSGADGAGGAGGA-----RLTNGGCG--GSGVGAAPTG--- 289
QY 196 NAYKKGYTDALSGMNGLSQLLNGGLGCG--GCGNAGTGLDSSLGKGLQNLSPV 252
DB 290 -----INGG-----GNGNGGMLTGDGSGGSLTGLGTSDDGTG----- 322
QY 253 DYQQLGNVGTGIGMRAQIALNDIGTHRSSTRSVYKCDRMAKEIGOFMDQYEVFG 312
DB 323 -----GNAAGFPGCGGNGGFFFGNGGDDGTGTLTNGNGD----- 358
QY 313 KPYQKPGQEVETDCKSMNAKSKPDDDCMTPASMEQFKAQMKIRPAGPTGN----- 368
DB 359 -----GNSVQTD-----GFL--RGHGDDGAGNAVGL 382

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2003, 08:46:41 ; Search time 11 Seconds

(Without alignments)
1519.542 Million cell updates/sec

Title: US-09-412-100-23

Perfect score: 2079
Sequence: 1 MSINTSGIGASTMOISIGGA.....DAMWAGDAINNMALGKLGAA 403

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	2050	98.6	403	1	HRPN_ERHAM
2	828.5	39.9	356	1	HRPN_ERMCA
3	718.5	34.6	463	1	HRPN_ERMCH
4	248	11.9	463	1	YA68_MYCRU
5	242	11.6	778	1	YQ34_MYCRU
6	241.5	11.6	957	1	Y278_MYCRU
7	238.5	11.5	515	1	Y140_MYCRU
8	238	11.4	1901	1	Y208_MYCRU
9	235.5	11.3	914	1	WA22_MYCRU
10	234.5	11.3	543	1	YP91_MYCRU
11	233.5	11.2	338	1	GRP_ARATH
12	232	11.2	801	1	Y747_MYCRU
13	229.5	11.0	603	1	YD25_MYCRU
14	229	11.0	491	1	YR98_MYCRU
15	214.5	10.3	384	1	GRP1_PERYH
16	213.5	10.3	481	1	LORI_MOUSE
17	211.5	10.2	747	1	SPD1_MOUSE
18	210	10.1	1156	1	GLH4_CAEEL
19	202.5	9.7	937	1	Y118_MYCRU
20	199	9.6	937	1	HRP1_CANAL
21	187.5	9.5	344	1	POPA_RALSO
22	186.5	9.5	465	1	GRP2_PHAVU
23	193	9.3	183	1	GRP2_ORYSA
24	191	9.2	252	1	GRP1_PHAVU
25	191	9.2	316	1	LORI_HUMAN
26	190	9.1	641	1	EBN1_EBV
27	186.5	9.0	1113	1	N116_YEAST
28	186	8.9	1387	1	TRDP_HUMAN
29	184.5	8.9	959	1	CSP_PLACB
30	178.5	8.6	419	1	N100_YEAST
31	177.5	8.5	1224	1	PER_DROME
32	175.5	8.4	734	1	YR2_CAEEL
33	174.5	8.4	672	1	PHX_MOUSE

34	174	8.4	212	1	EGG1_SCHUA	P19470 schistosoma
35	174	8.4	700	1	NOMA_DROME	O04047 drosophila
36	173.5	8.3	1380	1	DDX9_MOUSE	O70133 mus musculus
37	172	8.3	5263	1	FB0H_BOMMO	P05790 bombyx mori
38	171.5	8.2	1208	1	PER_DROVA	Q24767 drosophila
39	169	8.1	526	1	PUS_HUMAN	P35637 homo sapien
40	169	8.1	1596	1	NAM_DROME	P21519 drosophila
41	167	8.0	401	1	CSP_PLACB	P08674 plasmodium
42	166.5	8.0	393	1	CSP_PLACB	P14593 plasmodium
43	165.5	8.0	207	1	EGG2_SCHUA	P19469 schistosoma
44	164.5	7.9	404	1	CAZ_DROME	Q27294 drosophila
45	164	7.9	518	1	FUS_MOUSE	P56959 mus musculus

ALIGNMENTS

RESULT 1
HRPN_ERHAM STANDARD: PRT: 403 AA.

AC 001099:
DT 01-APR-1993 (Rel. 25, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Harpin (Harpin-EA).

GN HRPN.
OS Erwinia amylovora.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX Erwinia.

NCBI_TaxID=552;

SEQUENCE FROM N.A., AND SEQUENCE OF 1-15.

RC STRAIN-Ea321:

RX MEDLINE-92320301; PubMed-1621099;
RA Wei Z.-M., Lady R.J., Zumoff C.H., Bauer D.W., He S.Y., Collmer A.,
Beer S.V.;

RT "Harpin, elicitor of the hypersensitive response produced by the
RT plant pathogen Erwinia amylovora.";
RL Science 257:85-88(1992).

RN [2]

RP REVISIONS.

RC STRAIN-Ea321:
RA Lady R.J., Kim J.F., Beer S.V.;

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RT FUNCTION: ELICITS THE HYPERSENSITIVE RESPONSE (HR) IN THE PLANT
UPON INFECTION. HARPIN ELICITS HR IN NON-HOSTS AND IS ALSO
REQUIRED FOR PATHOGENICITY IN HOST PLANTS.

-1- SUBCELLULAR LOCATION: SECRETED; VIA THE HRP SECRETION PATHWAY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC EMBL: M92994; AAC31644.2; -

KM Hypersensitive response.

FT DOMAIN 1 270 GLY-RICH.

FT POLY-MET.

SQ SEQUENCE 403 AA: 39697 MW: 146FA642351DBE87 CRC64;

Query Match 98.6%; Score 2050; DB 1; Length 403;

Best Local Similarity 99.0%; Pred. No. 3e-115;

Matches 399; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MSINTSGIGASTMOISIGGAGNGGLGTSRONAGLGNSALGLGGGNDPTVNOIAGIL 60
DB 1 MSINTSGIGASTMOISIGGAGNGGLGTSRONAGLGNSALGLGGGNDPTVNOIAGIL 60
QY 61 TGMNDMMSMGGGGLMGGLGGGLGNGLGSGGLGEGLSNLDMDAGSLNTLTGSKGNN 120

Db 61 TGMAMMSMMGGGGLMGGLGGLGNGLGSGGLGEBLSNANLMDLGSINTLGSKGN 120

Qy 121 TTSTNSPLDQALGINSTSONDSTSGTSTSDSSDPMOQLKMFSEIMOSLFEGDQGT 180

Db 121 TTSTNSPLDQALGINSTSONDSTSGTSTSDSSDPMOQLKMFSEIMOSLFEGDQGT 180

Qy 181 OGSSSGKOPTEGEONAKKGYDALSGLMNGLSQLLGNGLGGGGGNAAGTGLDSSSL 240

Db 181 OGSSSGKOPTEGEONAKKGYDALSGLMNGLSQLLGNGLGGGGGNAAGTGLDSSSL 240

Qy 241 GKGGLQNLGSPVDYQQLGNAVGTIGMKAGIOLANDIGTHRSSTSFYNNKGDRAAKEI 300

Db 241 GKGGLQNLGSPVDYQQLGNAVGTIGMKAGIOLANDIGTHRSSTSFYNNKGDRAAKEI 300

Qy 301 GGFMDQPEYFVGKPOYKPGQEVKTDKSNAAKLSRDDGMPASMEQFNKAKIKR 360

Db 301 GGFMDQPEYFVGKPOYKPGQEVKTDKSNAAKLSRDDGMPASMEQFNKAKIKR 360

Qy 361 PMAGDTGNGMLQARGAGSSSLGIDAMMAGDAINNMALGKLGAA 403

Db 361 PMAGDTGNGMLQARGAGSSSLGIDAMMAGDAINNMALGKLGAA 403

RESULT 2

HRPN_ERMCA

ID HRPN_ERMCA STANDARD; PRT; 356 AA.

AC 047278; 356 AA.

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Harpin (Harpin-ECC) (Fragment).

GN HRPN.

OS Erwina carotovora.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Pectobacterium.

OX NCBI_TaxID=554;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-71;

RX MEDLINE-96405946; PubMed-8810071;

RA Cai Y., Madl L., Mukherjee A., Dumenyo C.K., Chatterjee A.K.;

RT "The RsmA-mutants of Erwina carotovora subsp. carotovora strain

RT Ecoc71 overexpress hrpNcc and elicit a hypersensitive reaction-like

RT response in tobacco leaves.";

RL Mol. Plant Microbe Interact. 9:565-573(1996).

CC -1- FUNCTION: ELICITS THE HYPERSENSITIVE RESPONSE (HR) IN THE PLANT

CC UPON INFECTION. HARPIN ELICITS HR IN NON-HOSTS AND IS ALSO

CC REQUIRED FOR PATHOGENICITY IN HOST PLANTS.

CC -1- SUBCELLULAR LOCATION: SECRETED; VIA THE HRP SECRETION PATHWAY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC EMBL; L/8834; AAB49733.1; -

CC Hypersensitive response.

CC FT DOMAIN 1 224 GLY-RICH.

CC FT NON TER 356 356

CC SQ SEQUENCE 356 AA; 108B46B9D27F9DE4 CRC64;

Query Match 39.9%; Score 828.5; DB 1; Length 356;

Best Local Similarity 48.2%; Pred. No. 9.4e-43;

Matches 196; Conservative 37; Mismatches 115; Indels 59; Gaps 11;

Qy 3 LNTSGAGASTMOISIGAGGNNGLL-GTSRONAGLGNSALGLGGNNDVTYNQALGLT 61

Db 2 LNTSGAGASTMOISIGAGGNNGLL-GTSRONAGLGNSALGLGGNNDVTYNQALGLT 61

Qy 62 GMMAMMSMMGGGGLMGGLGGLGNGLGSGGLGEBLSNANLMDLGSINTLGS 115

Db 57 TMMFMGSMGGG--MSGGLGGLGSSLGGLGGLG-GGLGGLGSSLGSSGSLG--CG 111

Qy 116 KGNNTSTNSPLDQALGINSTSONDSTSGTSTSDSSDPMOQLKMFSEIMOSLFEGD 175

Db 116 KGNNTSTNSPLDQALGINSTSONDSTSGTSTSDSSDPMOQLKMFSEIMOSLFEGD 175

Qy 112 LGC-----ALGAMNMANPSANMS-----LFSALEDLGG 143

Db 112 LGC-----ALGAMNMANPSANMS-----LFSALEDLGG 143

Qy 176 GDDTQSSSGKOPTEGEONAKKGYDALSGLMNGLSQLLGNGLGGGGGNAAGTGL 235

Db 176 GDDTQSSSGKOPTEGEONAKKGYDALSGLMNGLSQLLGNGLGGGGGNAAGTGL 235

Qy 144 GMSQGGGGLGNGKPPSPPEISATYQGVNDALSAILGNGLSQTKQ-----T 189

Db 144 GMSQGGGGLGNGKPPSPPEISATYQGVNDALSAILGNGLSQTKQ-----T 189

Qy 236 DGSLSGKGLQNLGSPVDYQQLGNAVGTIGMKAGIOLANDIGTHRSSTSFYNNKGDRA 295

Db 236 DGSLSGKGLQNLGSPVDYQQLGNAVGTIGMKAGIOLANDIGTHRSSTSFYNNKGDRA 295

Qy 190 SPQLGNNGLQGLSGAGAFNQLGSLTGMNVQKAGLDELNNISTHNDSPRYFYDKDRG 249

Db 190 SPQLGNNGLQGLSGAGAFNQLGSLTGMNVQKAGLDELNNISTHNDSPRYFYDKDRG 249

Qy 296 MAKEIGFMQYPEYFVGKPOYKPGQEVKTDKSNAAKLSRDDGMPASMEQFNKAK 355

Db 296 MAKEIGFMQYPEYFVGKPOYKPGQEVKTDKSNAAKLSRDDGMPASMEQFNKAK 355

Qy 250 MAKEIGFMQYPEYFVGKPOYKPGQEVKTDKSNAAKLSRDDGMPASMEQFNKAK 309

Db 250 MAKEIGFMQYPEYFVGKPOYKPGQEVKTDKSNAAKLSRDDGMPASMEQFNKAK 309

Qy 356 GMIKRPACDPTGNGMLQARGAGSSSLGIDAMMAGDAINNMALGKLGAA 402

Db 356 GMIKRPACDPTGNGMLQARGAGSSSLGIDAMMAGDAINNMALGKLGAA 402

Qy 310 GMIKSAIRGDPGTGNTLSARGNGASLGIDAMMAGDAINNMALGKLGAA 356

Db 310 GMIKSAIRGDPGTGNTLSARGNGASLGIDAMMAGDAINNMALGKLGAA 356

RESULT 3

HRPN_ERMCH

ID HRPN_ERMCH STANDARD; PRT; 340 AA.

AC 047278; 340 AA.

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Harpin (Harpin-ECH).

GN HRPN.

OS Erwina chrysanthemi.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Pectobacterium.

OX NCBI_TaxID=556;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-EC16 / AC1450;

RX MEDLINE-96172740; PubMed-8589405;

RA Bauer D.W., Wei Z.M., Beer S.V., Collmer A.;

RT "Erwinia chrysanthemi harpinch: an elicitor of the hypersensitive

RT response that contributes to soft-rot pathogenesis.";

RL Mol. Plant Microbe Interact. 8:484-491(1995).

CC -1- FUNCTION: ELICITS THE HYPERSENSITIVE RESPONSE (HR) IN THE PLANT

CC UPON INFECTION. HARPIN ELICITS HR IN NON-HOSTS AND IS ALSO

CC REQUIRED FOR PATHOGENICITY IN HOST PLANTS.

CC -1- SUBCELLULAR LOCATION: SECRETED; VIA THE HRP SECRETION PATHWAY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC EMBL; L39897; AAC31978.1; -

CC Hypersensitive response.

CC FT DOMAIN 1 203 GLY-RICH.

CC SQ SEQUENCE 340 AA; 34274 MW; EAD82C23731EF4C5 CRC64;

Query Match 34.6%; Score 718.5; DB 1; Length 340;

Best Local Similarity 42.6%; Pred. No. 3.1e-36;

Matches 172; Conservative 41; Mismatches 114; Indels 77; Gaps 10;

Qy 13 MOIST-GGAGGNNGLGTSRONAGLG-GNSA---LGLGGNNDVTYNQALGLTGMAMM 67

Db 1 MOITKHHIGDGLVSGIGAGGKLNKLNRAASSVSKLSTIDKLSALTSMA--- 57

Qy 68 SMAGGGGLMGGLGGLGNGLGSGGLGEBLSNANLMDLGSINTLGSKGNNTSTNS 127

ID	NAME	STANDARD	PRG	463	AA
Db	58	-----FCGALAGLIGAS-SKGLGMSNQLGGSFGN-----	GAQGSMTLSVPR	98	
Oy	128	PLDQALGINTSQNDSDSTSGTSDTSDDSPMOQLKRFSELMOSLFG-----	DGQDG	179	
Db	99	-----SGDALS-----	KMFDRALDILLGHDPVTKLTNSQ	130	
Oy	180	TOGSSSGCKOPTBGEONAYKKRGVYDALSGLMGNLSOLLGNGLGGGGGCGNAGTGLDSS	239		
Db	131	LANSMLNASQNTGNNNAFSGSVNNALSLTGLGSGSM-----	SGFSQPS	176	
Oy	240	LGCGKGLANTSGPVYQOLGNAVGTGIMKAGIQLANDIGTHERHSSTSFYKCDRAKAE	299		
Db	177	LGAGGLGGLGSGAGFNGLGNALIGCVQONALSLSVSHYDGNRNHPFKDKRGNAKE	236		
Oy	300	IGPFMDQYPEVFEKFPQYOKGGEQVKTDDKSMKALSKPDDGWTGPSMDFNKAQMIK	359		
Db	237	IGPFMDQYPELFEFGKPEYKDGKWSSPKTDKSMKALSKPDDGWTGPSMDFNKAQMIK	296		
Oy	360	RPMAGDTGNGTLAGAGGSSIGTIDAMAGDATNNALGRTGAA	403		
Db	297	SAVAGDTGNTNLNRGAGASLGDAAVAVBDKIANMSLGLANA	340		
RESULT 4					
YA68_MYCTU	ID	YA68_MYCTU	STANDARD	PRG	463
AC	053416				
DT	30-MAY-2000	(Rel. 39, Created)			
DT	30-MAY-2000	(Rel. 39, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Hypothetical PE-PGRS family protein Rv1068C.				
GN	Rv1068C OR MT1097 OR MTV017.21C.				
OS	Mycobacterium tuberculosis.				
OC	Bacteria; Actinobacteria (class); Actinobacteridae;				
OC	Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.				
OX	NCBI_Taxid=1773;				
RP	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=H37rv;				
RX	MEDLINE=98295987; PubMed=9634230;				
RA	Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,				
RA	Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,				
RA	Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,				
RA	Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,				
RA	Hornisy T., Jajelski K., Krogh A., McLean J., Moule S., Murphy L.,				
RA	Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,				
RA	Rutler S., Seeger K., Skelton S., Squares R., Squares R.,				
RA	Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;				
RT	"Deciphering the biology of Mycobacterium tuberculosis from the				
RT	complete genome sequence."				
RL	Nature 393:537-544(1998).				
RP	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=CDC 151 / Oshkosh;				
RA	Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,				
RA	Petersen J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,				
RA	Kolony J.F., Nelson W.C., Mayhew L.A., Emolova M.D., Salzberg S.L.,				
RA	Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,				
RA	Biswal W.;				
RT	"Whole genome comparison of Mycobacterium tuberculosis clinical and				
RT	laboratory strains"				
RL	Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.				
CC	-1- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGRS				
CC	SUBFAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				

DR	EHBL; AL021897; CAAT184.1;	-
DR	EHBL; AE006991; AAK45353.1; ALT_INTT.	
DR	TIGR; MT1097;	--
DR	tuberculin; RV1068c;	--
DR	InterPro; IPR000084; PE_region.	
DR	Pfam; PF00934; PE; 1.	
DR	Prodom; PD001223; PE_region; 1.	
RW	Hypothetical protein; Complete proteome.	
FT	CONFICT	218 218 T -> S (IN REF. 2).
FT	G- -> GGGAGGAGADGTGRKGDAGAGAGAGGWTGHGV	
FT	GDDGTGGCGGDXVGEPEPDGAAGAAGAG (IN REF.	
FT	CONFLICT	235 235
SQ	SEQUENCE	463 AA; 39305 MW; CF5696A7E9593952 CMC6A;
Query Match	11.9%; Score 248; DB 1; Length 463;	
Best Local Similarity	31.8%; Pred. No. 3.7e-08;	
Matches	93; Conservative 17; Mismatches 112; Indels 70; Gaps 13	
OY	8 LGASTM0ISTGAGAGNNGILSTSRNAGLIGGSALGLGCGNONPTVNAGLLTGMMMA 67	
Dd	203 IGAPVAGAGAGAGGAGTAGLFG---NGAGGAGAGAGAGAGCGD--GGSAGMTSGNGDA	
OY	68 SMAGGGGLMGGGGGLGAGGLGAGSGLGEGSLNALNDMLKSLNTLKSSKGNNTTSTYS 127	
Dd	257 GTGGGGGGMANGNGNSAGMLGNSGTGGC-----GGTAGA-GGGGNC-----NS 301	
OY	128 PLDQALGINSTQNDSITGSTDTSDSDPMOQLIKMFSEIMOSLFEDGDGDTGSS-- 184	
Dd	302 GIDPG--NCGGADGNAGNAGNGHGSA-----AKLFGDGAGAGAGMGST 344	
OY	185 -----SGGKQTEPEGDMNAKKGYMDALSGLMGNTLSQL-----GNGSIGGCOC--G 229	
Dd	345 GGTGGGGGREGGTGGGNGGHHGAGAGSGGTGTLGSGSGGTGGDNGGLGAGSGAKG 404	
OY	230 NAGTGLDSSISLGGLQNLSPDYVOLANAVTGIKKAG---IOALLNDIG 278	
Dd	405 NGGCGD----GGKG-----GDMLGNGGNGNGKGTGLMPGINTRG 445	
RESULT 5		
Y034_MXCTU	STANDARD; PRT; 778 AA.	
ID Y034_MYCTU		
AC P19933:		
DT 01-NOV-1997 (Rel. 35, Created)		
DT 01-NOV-1997 (Rel. 35, Last sequence update)		
DT 16-OCT-2001 (Rel. 40, Last annotation update)		
DE Hypothetical PE-PGRS family protein RV2634c.		
GN RV2634C OR WT2712 OR MTCY441.04C.		
OS Mycobacterium tuberculosis.		
OC Bacteria; Actinobacteria; Actinomycetia (class); Actinomycetidae;		
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.		
OX NCBI_Taxid=1773;		
RN [1]		
RP SEQUENCE FROM N.A.		
RC STRAIN-H37RV;		
RX MEDLINE=98295987; PubMed=9634230;		
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,		
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Telata F.,		
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,		
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,		
RA Hornsby T., Jagels K., Krogh A., McLaren J., Moule S., Murphy L.,		
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,		
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,		
RA Sultson J.E., Taylor K., Whitehead S., Barrell B.G.;		
RA "Deciphering the biology of Mycobacterium tuberculosis from the		
RT complete genome sequence."		
RL Nature 393:537-544(1998).		
RN [2]		
RP SEQUENCE FROM N.A.		
RC STRAIN-CDC 1551 / Oshkosh;		
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,		
RA Peterson J., Deboy R., Dodson R., Gwin M.L., Haft D., Hickey E.,		
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,		

[illegible]

QY 255 OOLGNVGTG-----IGMKAGIOALNDIGTHRSSTRSFVNKGDRAMAKEIGOFMDQYP 308
 Db 631 SLIIGSGGTGNGNSIGV-AGIGGAGRG-----GDAGL-----664
 QY 309 EVFKPKQYKQPGQEVKTDKSNAKALSKRPDDDKTPASMDQFNKAKMIRPMAGDTGN 368
 Db 665 -LFG---AACTGG-----HOAGGVPAVGAGGAGCN 690
 QY 369 GNLQARGAGSSIGIDAMAGDAINNALGKIG 401
 Db 691 GGLFANGAGAGAGGAFNA-AGGNGNGIGLFTGG 722

RESULT 7
 Y140_MYCTU STANDARD: PRT: 515 AA.
 AC 050594:
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical PE-PGRS family protein RV1840C.
 GN RV1840C OR MT1888 OR MTCY1A11.04 OR MTCY359.33.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID-1773;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV:
 RA MEDLINE-96295987; PubMed-9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Teale F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers R.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sultson J.E., Taylor K., Whitehead S., Barrett B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Kouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 [1-] SUBCELLULAR LOCATION: Integral membrane protein (potential).
 [1-] SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGRS
 SUBFAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: Z83859; CAB06114.1; -
 DR EMBL: AE007047; AAK46159.1; -
 DR TIGR: MT1888; -
 DR TubercuList: RV1840C; -
 DR InterPro: IPR000084; PE_region.
 DR Pfam: PF00934; PE.1.
 KW Hypothetical protein, Transmembrane, Complete proteome.
 FT TRANSMEM 1 21 POTENTIAL.

FT TRANSMEM 165 185 POTENTIAL.
 FT TRANSMEM 199 219 POTENTIAL.
 SQ SEQUENCE 515 AA; 43916 MW; 2EDC8B6C6D28B7F3 CRC64;
 Query Match 11.5%; Score 238.5; DB 1; Length 515;
 Best Local Similarity 27.3%; Pred. No. 1.5e-07;
 Matches 109; Conservative 23; Mismatches 128; Indels 139; Gaps 15;

QY 18 GGAGNNNGILGTRSRNAGIGNSALGIGGNGNDVNDVNLGTLTGMMAMMSMGGGGILMG 77
 Db 148 GGAGGDPAGLIGNG-GNGIGGPGATGLAG-----AGGVGLTGDDGNGGAGGILG 197
 QY 78 GGLGGGLGNGLGGSGGLGEGLSNALNDMLGSLNTLGSKGNNTTSTNPL-DOALGI 135
 Db 198 TGPGVATG-GIGGPGGAAYGLFPHGGAGAGGLGKAGAFGAGGAGTGCGLLYNGNGG 256
 QY 136 NSTSONDSTSGTSDTSDSDPMQOLLKMFSEIMOSLRGDGDTGSSSGGKOPTBEO 195
 Db 257 NVPSCAADGAGGDA-----RLIGNGGDG---GSVCAAPLG---289
 QY 196 NAYKKGYTDALSGINGNGLSQLLNGGIGG---GGGNNAGTGLDSSSLGKGLQNLSPV 252
 Db 290 -----LNG-----GNGNGMLYIGDGGSGGSLTGFSGDTG-----322
 QY 253 DYQULGNNAVGTGIGMKAGIOALNDIGTHRSSTRSFVNKGDRAMAKEIGOFMDQYPEYFG 312
 Db 323 -----GNMGFDDGGNGGFSFDDGGGCGGTGTLGNGD-----358
 QY 313 KFOYKQPGQEVKTDKSNAKALSKRPDDDKTPASMDQFNKAKMIRPMAGDTGN---368
 Db 359 -----GGSNVQTD-----GFL-RHGSDGNNAVGL 382
 QY 369 -GNLQARGAGSSIGIDAMAGDAINN---NALGKIG 402
 Db 383 IGMGAGGAGSAGTGFAPGGSGGNGGALLVNGGA 421

RESULT 8
 Y208_MYCTU STANDARD: PRT: 1901 AA.
 AC 053553;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical PE-PGRS family protein RV3508 precursor.
 GN RV3508 OR MTCV23.15.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID-1773;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV:
 RA MEDLINE-96295987; PubMed-9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Teale F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers R.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sultson J.E., Taylor K., Whitehead S., Barrett B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 [1-] SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGRS
 SUBFAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

OY 357 MIRPMAGTCGNGTQARGAGSSIGTIDAMAGDAINMALGKICA 402
 DB 551 W-----LFGNGGNGAGAGAATNGSAG-----AGGA-GGILFGTGA 586

RESULT 10

TP91_MTCU STANDARD; PRT; 543 AA.

AC 050630;
 DT 01-NOV-1997 (Rel. 35, Last Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Hypothetical PE-PGRS family protein Rv2591.
 GN Rv2591 OR MT2668.1 OR MTCY227.10C.
 OS Mycobacterium tuberculosis.
 CC Bacteria; Actinobacteria; Actinobacteridae; Actinobacteriales; Actinomycetales; Actinomycetaceae; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.

OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,

RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandram M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;

RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).

[2] SEQUENCE FROM N.A.

RC STRAIN-CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J.F., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;

RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGRS
 CC SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).

CC EMBL: Z77724; CAB01283.1;
 CC EMBL: AE007100; AAK46982.1; ALT_INIT.

CC TIGR: MT2668.1;
 CC Tuberculist: Rv2591;
 CC InterPro: IPR000084; PE_region.

CC DR PIR: P00934; PE; 1.
 CC DR ProDom: PD001223; PE_region; 1.
 CC KM Hypothetical protein; Complete proteome.

SO SEQUENCE 543 AA; 46287 MW; 597303395D2DF59 CRC64;
 Query Match 11.3%; Score 234.5; DB 1; Length 543;
 Best Local Similarity 28.0%; Pred. No. 2.8e-07;
 Matches 111; Conservative 31; Mismatches 121; Indels 133; Gaps 21;

OY 18 GGAGGNNGLGTSNONAGCGNSALGCGGNDNTVNOIAGLIGMMAMMSMAGC-----72
 DB 213 GGAGGADPLIGWG-GNGGPGGFAFNGGAGG-----GGAGSLEGGVAGGAGVSS 264

OY 73 ---GGIMG-GGILGIGGIGGSSGIGELISN-----ALNDMLGSLINTLASKGNNTT 122
 DB 265 EDVGTGAGGAGKRLFLGIGDGDAG--GTSNNNGDGGAGGTAGRLFLSLGGDGNCGA 323

OY 123 STNSPLDQALGINSTSONDSTSGTSDSDPMQQLMKFSEIMOSLFGDQDDGNOG 182
 DB 324 GT-----AIGSNM-----GGGAGGDS-----ALGTAAGSGSG 353

OY 183 SSSGKQPEEBQNNYKKTVDALSLGMLNGLSQLLNGGIGAGG--GNAAGTGLDSS 239
 DB 354 --LGG---FGEISTGDCGLGAGAYLLIGTGV---GGGIGGSGNGTGAGAGGTAT 404

OY 240 L-----GGKGLQNSGPDVVOQGNVCTGIGMAAGI-QALNDIGTRHSSTSFVKG 292
 DB 405 LIGLAGGGGGIGGFA-----VWNGNGV-GIGGGGGGAAIGIGAG-----GAG 449

OY 293 DRAMKEIQPFMDQYPEYFGKPYQKPGQEVKTDKSMARALSKPDDGWT PASMDFN 352
 DB 450 GAGGATVVG-----LGGNGDG-----G 467

OY 353 KAKGKIRPMAGD---TGNGMLQARGAGSSIGTIDA 385
 DB 468 DGGGLFSLGVDGDMAGNAGAMPANGGNGAGVIA 503

RESULT 11
 GRP_ARATH STANDARD; PRT; 338 AA.

AC P27483;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-AUG-1992 (Rel. 23, Last annotation update)
 DE Glycine-rich cell wall structural protein precursor.

OS Arabidopsis thaliana (mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC Eurosids II; Brassicales; Brassicaceae; Arabidopsids.

CC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-cv. Columbia;
 RX MEDLINE=92003708; PubMed=1912511;
 RA Quigley F., Villiot M.L., Mache R.;

RT "Nucleotide sequence and expression of a novel glycine-rich protein
 RT gene from Arabidopsis thaliana."
 RL Plant Mol. Biol. 17:949-952(1991).

CC -1- FUNCTION: RESPONSIBLE FOR PLASTICITY OF THE CELL WALL (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: Cell wall (Potential).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).

CC EMBL: X58338; CAA41249.1;
 CC PIR: S17732; KMMO.

CC DR Cell wall; structural protein; Repeat; signal.
 CC FT SIGNAL 20 POTENTIAL.
 CC FT CHAIN 21 338 GLYCINE-RICH CELL WALL STRUCTURAL
 CC FT DOMAIN 21 338 GLY-RICH.

SO SEQUENCE 338 AA; 23891 MW; 046A6B8C1A4E89EB CRC64;
 Query Match 11.2%; Score 233.5; DB 1; Length 338;
 Best Local Similarity 29.9%; Pred. No. 1.9e-07;
 Matches 80; Conservative 14; Mismatches 77; Indels 97; Gaps 10;

OY 7 GLGASTMOISIGGAGGNNGLGTSRONA--GIGGNSALGIGGNDNTVNOIAGLITGM 64
 DB 59 GLG-----GGAGGGGIGGAGGAGGAGGGLGAGGAGGLGAGH-----GGGIGG-- 100

[illegible]


```

DR   PIR: A26099; A26099.
KW   Cell wall; Structural protein; Repeat; Signal.
FT   SIGNAL      1      27
FT   CHAIN       28      384
FT   DOMAIN      41      384
FT   SEQUENCE    384 AA; 28777 MW; C0541C549417D18C CRC64;
SQ
Query Match      10.3%; Score 214.5; DB 1; Length 384;
Best Local Similarity 28.1%; Pred. NO. 3e-06;
Matches 77; Conservative 17; Mismatches 111; Indels 69; Gaps 9

QY  7 GLGASTMIOISIGAGGNNGLGTSRONGLGNSALGLGCGNQNDTVNOLAALITGMMH 66
Db  152 GFAGAG-----GGVGGAGAGGAGGVGGGGGFGGGGGGGGGGGGGGAGGAGG 201
QY  67 MSNMNGGGLMGCGGLGGLGNGLGSGGGGEGESLSNALNDMLDGGSLTLTSGKGNNTTSTN 126
Db  202 ---GAGGGGLGAGGGGGGGGGGGGGGGGGGGGGSGH-----GGGFGAGGGVGG----- 243
QY  127 SPLDQALGINSTONSNDSTSGTDSSTDSDDPQQLMKFSEIMQSLFGKGD-----GTQ 181
Db  244 -----GVGGGAAGGGGGGGGGGGGGGGG-----LGGGSHGGGFGAG 280
QY  182 GSSSGGKOPTGEEDNAYKKVTDALSGLMGNLSOLLGNGGLGGGGGAGTGL--DGSS 239
Db  281 GGVGGGAAGVGGGGGGGGGGGGGGGGGGGGH-----GGFGAGGVGGGAGGGLGAGGGGA 335
QY  240 LGGKQLNLSGPVDYQQLGNNAVGTGIRKAGGAIQA 273
Db  336 GGGGGIGGGHGG-----GFGVVGIGIGVGVCA 363

```


Db 110 -TTGAGS-----LGLDPTQTGDDSLSGAGQTSMS-PMQMLKIPADITQSLFGD-QDG 161
 QY 180 TCGSSSGGKOPTGEGEONAYKGYTDALSGIMNGLSQLLNGGLGGGQGNAGTGLDSS 239
 Db 162 ASGGA-GRQPSQDEONAYKGYTDALTAEMGGGLSQVAGNSEGLDGG-MGLG-GGNG 218
 QY 240 LGGKGLQNTSGPVDYQOLGNNAVGTGIGMKGIALNDIGHRSSSTRSPFNKGRAMAKE 239
 Db 219 LGGKGLQNTSGPVDYQOLGNNAVGTGIGMKGIALNDIGHRSSSTRSPFNKGRAMAKE 239
 QY 300 ICGFMQYEVFEKGPYOKGPEVTDKSNAPKALSKPDDGMPASMEQFNKAKGMIK 359
 Db 279 VQGFMDQYEVFEKGPYOKGPEVTDKSNAPKALSKPDDGMPASMEQFNKAKGMIK 359
 QY 360 RPAAGDTGNGNLQARGAGSSSLGIDAMAGDAINNALGKIGAA 403
 Db 339 SAMAGDNGNINLQARGAGSSSLGIDAMAGDAINNALGKIGAA 382

RESULT 2

QY 09KH45 PRELIMINARY; PRT; 365 AA.
 AC 09KH45;
 DT 01-OCT-2000 (T-EMBLrel. 15, Created)
 DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
 DE Harpin.
 GN HRPN.
 OS Pantoea agglomerans pv. gypsophillae.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Pantoea.
 OX NCBI_TaxID=48984;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-21171042; PubMed-11277443;
 RA Mor H., Manulis S., Zuck M., Nizan R., Coplin D.L., Barash I.;
 RT Genetic organization of the hrp gene cluster and *hrp* operon in
 RT *Erwinia herbicola* pv. *gypsophillae*.
 RL Mol. Plant Microbe Interact. 14:431-436(2001).
 DR EMBL: AF271716; AAF76342.1;
 SQ SEQUENCE 365 AA; 36784 MW; 6F79BBDD95AC2A CRC64;

Query Match 48.2%; Score 1002.5; DB 2; Length 365;
 Best Local Similarity 57.9%; Pred. No. 3.4e-50;

Matches 223; Conservative 40; Mismatches 83; Indels 39; Gaps 14;

QY 1 MSINTSGTGAFTQISIGGAGGNNGLGTSRONAGLGSNSALGLGGGNDYVNOIAGL 60
 Db 1 MSINTPLGIPAMQISL---GDSNGLPGRMLQNAVLN-----SKGDSIDETINOLAGAL 52
 QY 61 TGMAMAMSMGCGGLMGGLGGGLGNGLGG--SGGLAEGISNALNDMGAGSLNT-LGSKG 117
 Db 53 TAMAMAMAMGSGGL--GGGLGTRA--GGELAGQSGGALGGPLYGNGGQGLGALGSLGTCG 110
 QY 118 GNNITSTNSPLDQALGINSTSGNDSTSGTSTSS-----DPMQILLKMSSEIMOSL 172
 Db 111 G-----ASSSLDKALDL-----DPTLDGGGLSNGNGTYSPIIDQILKMALEWQSM 157
 QY 173 FEDGQDGTGSSSGGKOPTGEGEONAYKGYTDALSGIMNGLSQLLNGGLGGGQGNAG 232
 Db 158 FG-GQGDITSGRGF-GNQPTPDQONATKVTYDALTYVMGGGLSQMGSS-CTGGGANGSIG 214
 QY 233 TGLDSSSLGKGLQNTSGPVDYQOLGNNAVGTGIGMKGIALNDIGHRSSSTRSPFNK 232
 Db 215 L---GNGLAGKGLQELNAGPADYEQLGVCYGTGVMAGIETALNNIGTESDSTRFTINKD 271
 QY 293 DRAMAEIGOPMDQYEVFEKGPYOKGPEVTDKSNAPKALSKPDDGMPASMEQFN 352
 Db 272 DMSLAIEVQGFMDQYEVFEKGPYOKGPEVTDKSNAPKALSKPDDGMPASMEQFN 331
 QY 353 KAKGMIKRPAGDTGNGNLQARGAG 377
 Db 332 KAKGMIKSAEGDT--GNLQATGEG 354

RESULT 3

QY 093FY8 PRELIMINARY; PRT; 356 AA.
 AC 093FY8;
 DT 01-DEC-2001 (T-EMBLrel. 19, Created)
 DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)
 DE HRPN.
 GN HRPN.
 OS Pectobacterium carotovorum subsp. carotovorum.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Pectobacterium.
 OX NCBI_TaxID=555;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SCC1;
 RA Rantakari A., Viitaharju O., Vaehaemiko S., Taira S., Palva T.E.,
 RA Saarilahti H.T., Romantschuk M.;
 RT "Type III secretion contributes to the pathogenesis of soft-rotting
 RT plant pathogen *Erwinia carotovora*: Partial characterization of the hrp
 RT gene cluster."
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF302656; AAK97283.1;
 SQ SEQUENCE 356 AA; 35534 MW; 154C9D885810DCB8 CRC64;

Query Match 40.4%; Score 840.5; DB 2; Length 356;
 Best Local Similarity 48.6%; Pred. No. 6.4e-41;

Matches 198; Conservative 36; Mismatches 114; Indels 59; Gaps 11;

QY 3 LNTSGTGAFTQISIGGAGGNNGLGTSRONAGLGSNSALGLGGGNDYVNOIAGL 61
 Db 2 LNTSGTGAFTQISIGGAGGNNGLGTSRONAGLGSNSALGLGGGNDYVNOIAGL 56
 QY 62 GMMAMAMSMGCGGLMGGLGGGLGNGLGG--SGGLAEGISNALNDMGAGSLNT-LGSKG 115
 Db 57 TMMFMGSMGCG--MSGGLGGLGSLGGLGGLG--GG 111
 QY 116 KGNNTTSTNSPLDQALGINSTSGNDSTSGTSTSS-----DPMQILLKMSSEIMOSL 175
 Db 112 LGG-----ALGAGNANMNSAMGS-----LTFSLIEDLGG 143
 QY 176 GQDGTGSSSGGKOPTGEGEONAYKGYTDALSGIMNGLSQLLNGGLGGGQGNAGTGL 235
 Db 144 GMSQDQGGGLFNGKQPSPEISATYGTGVDLSAILNGLSQTRGQ-----T 189
 QY 236 DGSLLGKGLQNTSGPVDYQOLGNNAVGTGIGMKGIALNDIGHRSSSTRSPFNK 235
 Db 190 SPLQGNNGIGLGSAGAFNQGLSTGLMSYGAQLQELNNISTHNDSPRYFYKEDRG 249
 QY 296 MAKEIGOPMDQYEVFEKGPYOKGPEVTDKSNAPKALSKPDDGMPASMEQFNKAK 355
 Db 250 MAKEIGOPMDQYEVFEKGPYOKGPEVTDKSNAPKALSKPDDGMPASMEQFNKAK 309
 QY 356 GMIKRPAGDTGNGNLQARGAGSSSLGIDAMAGDAINNALGKIGAA 402
 Db 310 GMITNVAAGDTGNTLSARGNGASLIGIDAMIGDRIVNNGKLKISS 356

RESULT 4

QY 09EXP0 PRELIMINARY; PRT; 344 AA.
 AC 09EXP0;
 DT 01-MAR-2001 (T-EMBLrel. 16, Created)
 DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (T-EMBLrel. 16, Last annotation update)
 DE Harpin protein.
 GN HRPN.
 OS *Erwinia chrysanthemi*.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Pectobacterium.
 OX NCBI_TaxID=556;

OY 257 LGNAVGTGIGMKAGTALNDIGTHRSSTRFVNKDRAMAKEIGOPMDQYPRVGRKPOY 316
 DB 1214 -GNA---GIG-----GDGSGGNGCG-GDSGSGIGGQPGF 1243
 OY 317 OKPGQGEVETDDKSNKALSKPDDDMTPASMEQFKAKGIMKRPVAGPTGNGNLQARGA 376
 DB 1244 AGGRPG-----KGGAGNAGTGTGNGSGA 1267
 OY 377 GGSISGIDAMMAGDAINNALGKLG 401
 DB 1268 GGAGGCGAGAGAGISFSGNSNGGTG 1292

RESULT 7

OY 007224 PRELIMINARY; PRT; 591 AA.
 AC 007224;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical protein RV0297.
 GN RV0297 OR MTCY63.02.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID-1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RX MEDLINE-98295987; PubMed-9634230;
 RA Gordon S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Cole S.T., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornby T., Jagsels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sultston J.E., Taylor K., Whitehead S., Barrett B.G.,
 RA "Deciphering the Biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 DR EMBL: Z96800; CAB09596.1; -
 DR Tuberculist: RV0297; -
 DR InterPro: IPR000084; PE_region.
 DR Pfam: PF00934; PE_1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 591 AA; 49058 MW; 4E4675ED4297F5A CRC64;

Query Match 12.3%; Score 255; DB 16; Length 591;
 Best Local Similarity 26.8%; Pred. No. 4e-07;

Matches 123; Conservative 24; Mismatches 168; Indels 144; Gaps 20;

OY 7 GIGASTMQISIGGAGNNGILGTSRONAGIGNSALGIGGONNDVYNOLAGILTGMMMM 66
 DB 141 GNGGSGAPGQAGAGGATGFFG-----NGNGGDCGAGANGC-----AGGTAGMFFG 187
 OY 67 MSMAGGGGLMG-----GIGCGGIGLN-----GIGSGGIGLEGISNALNDML----- 106
 DB 188 FGNGGAGGIGVAGINGIGAGGAGCGNAGFGNGNGGAGGAGAGVAVNAGLATPVPY 247
 OY 107 -----GSLNTLGSKG--GNMTSTNSPLDQALGINSTQNDSTSG-----TDSSTDS 154
 DB 248 PAANGNGNLVGVPTAGGADGANGSALGQAGAGCGDGNASTSGIGTIAOTGAGGA 307
 OY 155 SDPMQOLLKMFSEIMSLFQDGDGTGSSSGKOPTEBQNAKKGVTDALSGLNGNL 214
 DB 308 G-----GAGDGAAPGNGNGSGVEHT-----GATGS-SASGNGA 342
 OY 215 SOLLNGGIGL-----GGGNGACTGLDSSLSGKGLQNTSGPV 252
 DB 343 TG--GNGGVGAPGAGGNGGNGHVSNGSVNTAGAGKGGNGGTGAGGP--GGHGGSVLSGPV 399
 OY 253 DYQOGLNAVGTGIGMKAGTALNDIGTHRSSTRS-----FVNKDRAMAKEIGOPMDQY 308

DB 400 GDSGNGAGAGD---GAGVSATDIAGTGGRGNGGHEGLWIGNGCGAGAGV----- 449
 OY 309 EYFGKPYQYKPGQGEVETDDKSNKALSKPDDDMTPASMEQFKAKGIMKRPVAGPTGNGNLQARGA 376
 DB 450 -----GVC-----GAGAGALGHHGGG-----GSYNTPIGGSSEAG 480
 OY 365 DTGNGNLQARGAGSSSLGIDAMMAGDAINNALGKLGAA 403
 DB 481 DGGKGGIGDGGGRGIFG--QPGAGAGAGAGVGGAGAGA 517

RESULT 8

OY 08VKN3 PRELIMINARY; PRT; 622 AA.
 AC 08VKN3;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE PE_PGRS family protein.
 GN MT0311.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID-1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CDC 1551 / OSHKOSH;
 RA Rietschmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwin M.L., Halt D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Uitterback T., Weidman J., Khoult H., Gill J., Mikula A.,
 RA Bishop W.,
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE006938; AAK44534.1; -
 DR TIGR: MT0311; -
 DR InterPro: IPR00104; Antifreeze_1.
 DR Pfam: PF00934; PE_1.
 DR PRINTS: PR00308; ANTIFREEZE1.
 SQ SEQUENCE 622 AA; 52747 MW; B27AD977A4DE8B67 CRC64;

Query Match 12.2%; Score 254; DB 16; Length 622;
 Best Local Similarity 27.5%; Pred. No. 4.8e-07;

Matches 125; Conservative 22; Mismatches 172; Indels 136; Gaps 20;

OY 7 GIGASTMQISIGGAGNNGILGTSRONAGIGNSALGIGGONNDVYNOLAGILTGMMMM 66
 DB 172 GNGGSGAPGQAGAGGATGFFG-----NGNGGDCGAGANGC-----AGGTAGMFFG 218
 OY 67 MSMAGGGGLMG-----GIGCGGIGLN-----GIGSGGIGLEGISNALNDML----- 106
 DB 219 FGNGGAGGIGVAGINGIGAGGAGCGNAGFGNGNGGAGGAGAGVAVNAGLATPVPY 278
 OY 107 -----GSLNTLGSKG--GNMTSTNSPLDQALGINSTQNDSTSG-----TDSSTDS 154
 DB 279 PAANGNGNLVGVPTAGGADGANGSALGQAGAGCGDGNASTSGIGTIAOTGAGGA 338
 OY 155 SDPMQOLLKMFSEIMSLFQDGDGTGSSSGKOPTEBQNAKKGVTDALSGLNGNL 214
 DB 339 G-----GAGDGAAPGNGNGSGVEHT-----GATGS-SASGNGA 373
 OY 215 SOLLNGGIGL-----GGGNGACTGLDSSLSGKGLQNTSGPV 252
 DB 374 TG--GNGGVGAPGAGGNGGNGHVSNGSVNTAGAGKGGNGGTGAGGP--GGHGGSVLSGPV 430
 OY 253 DYQOGLNAVGTGIGMKAGTALNDIGTHRSSTRS-----FVNKDRAMAKEIGOPMDQY 312
 DB 431 GDSGNGAGAGD---GAGVSATDIAGTGGRGNGGHEGLWIGNGCGAGAGV----- 473
 OY 313 KPOYOKPGQGEVETDDKSNKALSKPDDDMTPASMEQFKAKGIMKRPVAGPTGNGNLQARGA 376

Db 474 -----GGVGGVGGVGGAGAGAGATGGGGD-----GSVMTPIGSGEAGDGGK 515
QY 369 GNLQAGAGAGSSIGIDAMMAGDALNNMALGKLGAA 403
Db 516 GGLGGGGGGGGTGG--QFGAGGAGAGAGVGGAGCA 548

RESULT 9

Q8VK71 PRELIMINARY; PRT; 628 AA.
ID 08VK71
AC 08VK71
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE PE_PGRS family protein.
GN MT3476.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN= CDC 1551 / OSHKOSH;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwin M.L., Hatt D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT *Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: A5007154; AAK47814.1;
DR TIGR: MT3476;
DR InterPro: IPR000104; Antifreeze_1.
DR InterPro: IPR001343; Hemlysn_Ca_bind.
DR InterPro: IPR000084; PE_region.
DR InterPro: IPR002173; PfkB.
DR Pfam: PF00934; PE; 1.
DR PRINTS: PR00308; ANTIFREEZE1.
DR PRINTS: PR00313; CABINDGRPT.
DR ProDom: PD001223; PE_region; 1.
DR PROSITE: PS00583; PFKB_KINASES_1; UNKNOWN_1.
SQ SEQUENCE 628 AA; 53828 MW; 9F62B1E7CC2C567E CRC64;

Query Match 12.1%; Score 251.5; DB 16; Length 628;
Best local Similarity 30.6%; Pred. No. 6.8e-07;
Matches 118; Conservative 16; Mismatches 155; Indels 97; Gaps 21;

QY 18 GGAGG--NNGLLGTSRONAGLGGNSALGLGGNONDTVNLGLLTGMMMMMSMGGGGL 75
Db 320 GGAGGANGNRPATVSTANAGAGNGANG--GDGANGAGAGAGGAAAGSSVGDDGNGG 378
QY 76 MCG-----GLGGGCGNGIGSGGGLGEG--SNALNDMLGSLNTLASKGGNNTTSTTNSPLD 130
Db 379 AGGTGTNGHAGGAG--GAGGAGGGRGMLVYGNGNGNGNG--GAAGNGCAIGCTG-- 430
QY 131 OALGINSTGNDSTGSDTSDSSDPMQLLKMFSLMSTLFGDGDGTGTGSSSGGKOP 190
Db 431 -AGGV--PANGGNSALGT-----QPVGGDGGGGNGGTG-- 463
QY 191 TGEONAYKKGVTDALSG--LMGNCISQLLNGGIG--GGGGNAGTGLDSSSLGKGLON 247
Db 464 TGRGGRGGGGGAGAGASGMLMGNGNG--GNGGTGGSGGAGGNGIGDDAG--GGNATST 520
QY 248 LSGPVDTQQLGNAVGTGIGKAGIQAINDIGTRHRSSTRSFVNGDRAMAKELIQFMDQY 307
Db 521 SSTPFDVHG--GNG--GAGGDAG-----HGTTGGDGGDGGHAGTGGRGGL-- 562
QY 308 PEYFGKPYQKGGQEVKTKDMSAKALSKRPDDGKTPASMEQNKAKGATKRMAGDTG 367
Db 563 -----AGQHRANSNGGGGGGTGAGAGTGG--TPGSGN-----AGTGTG 596

QY 368 GNLQAGAGAGSSIGIDAMMAGDALN 393
Db 597 TGNADSTNGPGSDG-----LGDAFN 618

RESULT 10

Q8VK71 PRELIMINARY; PRT; 635 AA.
ID 08VK71
AC 08VK71
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE PE_PGRS family protein.
GN MT1123.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN= CDC 1551 / OSHKOSH;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwin M.L., Hatt D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT *Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: A5006992; AAK45381.1;
DR TIGR: MT1123;
DR InterPro: IPR000084; PE_region.
DR Pfam: PF00934; PE; 1.
DR ProDom: PD001223; PE_region; 1.
SQ SEQUENCE 635 AA; 51726 MW; A461A61100C7C5D4 CRC64;

Query Match 12.0%; Score 250; DB 16; Length 635;
Best local Similarity 27.2%; Pred. No. 8.3e-07;
Matches 120; Conservative 23; Mismatches 152; Indels 146; Gaps 20;

QY 7 GLGASTMGIISGGAGGNGLLGT---SRONAGLGGNSALGLGGNONDTVNLGLLTGM 63
Db 293 GSGGHALLMGAGGAGGNGSGCTGAGAGSTGAGAGNGAGCGGCT-----GGLIFG- 343
QY 64 MAMMSMGGGGGIMG-----GGL-----GGGIG--NGIGSGSGGLGEGLSNLMIMG 107
Db 344 -----NGAGAGCATAGAGAGANGVSTTNGGTGNGGIGGTGSSGAGGNH--GILG 395
QY 108 -----GSLNTLGSKGGNNTTSTTNSPLDQALGINSTGNDSTGSDTSDSSDPMQ 160
Db 396 VGGAGGAGASGADDRGAGGTGTTSS--DGGAGGDDGGDNGAGAGTGG-- 444
QY 161 LKMFSEIMOSLFGDGDGTGTGSSSGGKOPTGSDNAYKKGVTDALSGLMGNG----- 213
Db 445 -----LFGAGNGNGGPG--GSGGAADIGGNGAGNGGCTGDNNGNGSGSGGAGSGG 492
QY 214 -----LSQLLNGGL--GGGGGN--AGTGLDSSSLGKGLONLSGPDVQQLGNAV 262
Db 493 DGGAGGNGAALFNGAGGAGGGRGNGAGAGGLGGGFFGLPGL--NGSG-----GGG 544
QY 263 TGIGMKAGIQALNDIGTRHRSSTRSFVNGDRAMAKELIQFMDQYPEYFGKPYQKPGQ 322
Db 545 GGNKAPGCVLYLNGN-----GAGGQSSGGIG-----GPG- 573
QY 323 EVKTDKSMKALSKRPDDGKTPASMDQFNKAKGIRPNAGDYGNGNTQARGAGSSSLG 382
Db 574 -----ATGAGAGKGGDGG-----DAQLGGDGGNGNGAGAGTGTGTPG 610
QY 383 IDAMMAGDALNNMALGKLGAA 403
Db 611 PGDPGGSGGLGLLFGQTGTA 631

RESULT 11
 ID 08VJ15 PRELIMINARY; PRT; 775 AA.
 AC 08VJ15;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE PE_PCRS family protein.
 GN MT3495.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 RX NCBI_TaxID-1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CDC_1551 / OSHKOSH;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwin M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Esmolaeva M.D., Salzberg S.L.,
 RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE007155; AAK47832.1; -
 DR TIGR; MT3495; -
 DR InterPro; IPR000084; PE_region.
 DR InterPro; IPR002173; PfKB.
 DR Pfam; PF00934; PE; 1.
 DR ProDom; PD001223; PE_region; 1.
 DR PROSITE; PS00583; PFKB_KINASES_1; UNKNOWN_2.
 SQ SEQUENCE 775 AA; 65448 MW; 5256A18546BDD53 CRC64;
 Query Match 12.0%; Score 250; DB 16; Length 775;
 Best Local Similarity 28.7%; Pred. No. 1e-06;
 Matches 96; Conservative 23; Mismatches 101; Indels 114; Gaps 15;
 QY 2 SLNTSGASTMIOISIGAGGNNGLTGRQ-----NAGLGGN-SALGIG 45
 DB 446 SAGTGGVAGS-----GGTGGNAGLIGAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAG 499
 QY 46 G---GNQNDTVNQLAGLITGMAMMMMSMMGGGGLMGGLG-GGLG-----NGLGSGG 93
 DB 500 GGLYGGGAGGAGGAGG-----GANIAGNGSDGAGAGHGGAGSARLIGAGRGKGGG 552
 QY 94 LEBGLSNALNDMLGSLNTLGRGN-----NTSTNSPLD 130
 DB 553 AGGNTAGRRADAIAGT-----GGDGGNGGNGGLLSGNAGAGHGGAGGAGSSSTATTCTPT 608
 QY 131 QALGINTSONDSTSGTSTSDSSPMQQLKMFSEIMOSLFEGDQDDTGGSSSGGKOP 190
 DB 609 GATGGGNGGAGAGGTGFTGSG-----GIGGNGAGGTGGGAG----- 646
 QY 191 TEGDONAYRKGYTDAL-----SGIMNGLSQLLNGGLG--GGGAGNAGTGLDSSSLGK 243
 DB 647 -----VALSVGSGTGGGNGSGGGLGCGGSGSLFNGNGAGAGVATGNGSGSGIGPASYG 701
 QY 244 G-----LQNLGSPVDYQQLGNVGTGIGKAG 270
 DB 702 GSKGGAAGAGLGG-----QIGNGSGSGSGGAGG 730
 RESULT 12
 ID 08VJ19 PRELIMINARY; PRT; 1217 AA.
 AC 08VJ19;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE PE_PCRS family protein.
 GN MT3615.3.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.

OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 RX NCBI_TaxID-1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CDC_1551 / OSHKOSH;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwin M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Esmolaeva M.D., Salzberg S.L.,
 RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE007155; AAK47974.1; -
 DR TIGR; MT3615; -
 DR InterPro; IPR000084; PE_region.
 DR InterPro; IPR002173; PfKB.
 DR Pfam; PF00934; PE; 1.
 DR ProDom; PD001223; PE_region; 1.
 DR PROSITE; PS00583; PFKB_KINASES_1; UNKNOWN_3.
 SQ SEQUENCE 1217 AA; 94463 MW; 2F5BDC5A30A056E CRC64;
 Query Match 12.0%; Score 249.5; DB 16; Length 1217;
 Best Local Similarity 25.5%; Pred. No. 1.8e-06;
 Matches 112; Conservative 27; Mismatches 148; Indels 153; Gaps 17;
 QY 4 NTSGLASTMIOISIGAG-----CGNGLTGRNAGLGSNALS-----LGG 46
 DB 731 NPTGIGGTGGTGGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 789
 QY 47 GNQNDTVNQLAGLITGMAMMMMSMMGGGGLMGGLG-----GGLG--NGLGSGGIGEG 97
 DB 790 GADADPGATG--TGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 847
 QY 98 LSNALNDMLGGS--LNTLGRGNNTSTNSPLDQALGINSQNDSTSGTSTSDSS 155
 DB 848 FSRGSGGTGGTGGVGTGGDGN-----AGGADDPKGGGTGGTGGSG 894
 QY 156 DPMQQLKMFSEIMOSLFEGDQDDTGGSS-----SGGKOPTEGONAYRKGYTDALSG 208
 DB 895 -----GAGSGGAGNNGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 934
 QY 209 LKNGLSQLLNGGLG-----GGGAGNAGTGLDSSSLGKGLQLNLGSPVDYQQLGNAY 261
 DB 935 TGGTG-----GTGGKGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 983
 QY 262 GTGIGKAGIQLANDIGTHHSSTRFVFNKGDPAKKEIGQFPDQIPEYFGKFOYOKRG 321
 DB 984 --GIG-----GDSGGGNGGQ--GDSGSGLGQDGFAGGPR 1015
 QY 322 QEVKTDKSWAKALSRPDDGKTPASWQFNKAKIKRPMAGDTGNGNLQARGAGSSSL 381
 DB 1016 G-----KGGAGGNAGTGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1039
 QY 382 GIDAMAGALNNMALGKIG 401
 DB 1040 QGAGAGAGISFSNGSGTG 1059
 RESULT 13
 ID 050415 PRELIMINARY; PRT; 731 AA.
 AC 050415;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE PCRS-family protein.
 GN RV3368 OR MTV004.46.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 RX NCBI_TaxID-1773;
 RN [1]

